From Department of Cell and Molecular Biology
Karolinska Institutet, Stockholm, Sweden

DEVELOPMENTAL DYNAMICS OF CARDIAC PROGENITORS AND THEIR ROLE IN CONGENITAL HEART DEFECTS

Mimmi Mononen

Karolinska Institutet

Stockholm 2021

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021

ISBN 978-91-8016-158-9
Cover illustration by Iida Valkonen

Developmental dynamics of cardiac progenitors and their role in congenital heart defects

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Mimmi Mononen

The thesis will be defended in public at Biomedicum, Karolinska Institutet, Stockholm, Sweden, on Tuesday the 23rd of March, 2021

Principal Supervisor:
Prof. Kenneth Chien
Karolinska Institutet
Department of Cell and Molecular Biology

Co-supervisor(s):
Asst. Prof. Fredrik Lanner
Karolinska Institutet
Department of Clinical Science, Intervention and Technology

Dr. Christian Riedel
Karolinska Institutet
Department of Biosciences and Nutrition

Dr. Chuen Yan Leung
Karolinska Institutet
Department of Medicine, Huddinge

Opponent:
Prof. Christine Mummery
Leiden University
Department of Anatomy and Embryology

Examination Board:
Prof. Martina Brueckner
Yale University
Department of Pediatric Cardiology

Prof. Eva Hellström Lindberg
Karolinska Institutet
Department of Medicine, Huddinge

Dr. Igor Adameyko
Karolinska Institutet
Department of Physiology and Pharmacology

To my family

"The heart has its reasons which reason knows not."
— Blaise Pascal

POPULAR SCIENCE SUMMARY OF THE THESIS

Congenital heart defect, a heart condition present at birth, is diagnosed in one in every 100 newborn babies. Congenital heart defects can affect different parts of the heart, and the clinical symptoms seen in a patient depend on the type of the defect. Some specific complex heart malformations cause cyanosis, which appears as a bluish skin color due to low oxygen content of the blood. In these patients, the oxygen poor blood which should be directed to the lungs for reoxygenation passes to the "wrong" side of the heart through a hole between the heart chambers, and enters the blood circulation through the aorta. Additional malformations are often present, which further hamper the heart's pumping function and the blood flow through the heart. The symptoms range from none to severe, and in the latter case a surgery or other procedures are needed soon after birth to ensure the baby's survival.

The cause of the heart malformation found at birth is often unknown. Some risk factors have been identified, including maternal consumption of certain medications and drugs, alcohol, tobacco and maternal obesity. Children of parents with heart defects are also at risk, and genetics has been proposed as the main cause of cardiac birth defects. The majority of congenital heart defects are not associated with other diseases, but they are sometimes seen in patients with Down syndrome or other genetic syndromes and rare diseases.

The congenital heart defects that have a genetic cause often arise early in the embryonic and fetal development, when something goes wrong on the molecular and cellular level. In some cases, the heart defect can be detected during pregnancy. Soon after birth, corrective surgeries can be carried out in order to close holes between cardiac chambers and redirect the abnormal blood flow. Although surgeries have significantly increased the lifespan and quality of life of the patients, congenital heart defects remain a major cause of child death globally. Therefore, development of new treatment strategies is a major goal of heart research. In order to develop novel therapies and improve the existing ones, more needs to be known about the genetic factors and disease mechanisms taking place during the early heart development before the baby is born. This kind of knowledge is the first step toward personalized gene and cell therapies and earlier diagnostics.

To access and study human embryos in the earliest steps of heart development, when many of the congenital defects arise, is currently one of the great challenges of cardiovascular research. Therefore, mouse embryos and other animals have been traditionally used instead to study the genetics, cells and tissues related to heart development and disease. In the past decades, researchers have identified specific groups of cells and cellular mechanisms that are likely involved in congenital heart defects. The early embryonic cells, which later build the heart, are called cardiac progenitors. In the research included in this doctoral thesis, we aimed to increase the understanding on *human* heart progenitors in early heart development and their various roles in congenital heart defects.

To access and model early human heart development, which cannot be accessed in human embryos, we used a so-called *in vitro* differentiation system. In this approach, human

embryonic stem cells are converted into cardiac muscle cells, or cardiomyocytes, in a laboratory setting. Interestingly, although lacking the 3D structure of the developing heart, the heart cells in a cell culture dish follow similar developmental patterns as their counterparts in the developing human embryo, and actually start beating around day ten of the experiment. Due to the technical advancements, the differentiation process from a stem cell to a cardiomyocyte can be studied at the level of individual cells. This is what we did in the first study presented in this thesis in order to fully understand the road from a stem cell to cardiomyocyte. Using computer algorithms, we could visualize this path, and found branch points of different cell fates diverging from the same starting cells. Our study revealed the similarities between human embryonic heart and the heart development in a dish, identified a cell type related to the development of the connective tissue of the heart, as well as the developmental path of heart progenitors.

In the second study we wanted to study the genetic causes of a specific type of complex cyanotic congenital heart defect in patients with healthy parents. To this end, parents of small children carrying this kind of heart defect were recruited to the study, and the genetic material of both the parents and the child was extracted and decoded computationally. This allowed us to identify genetic mutations that arose in the early development or in the reproductive (sperm and egg) cells leading to the child but not the parents having a heart defect. In some cases, both parents were carriers of risk mutations, which in combination caused a heart defect in the child. Bringing these observations together, we identified groups of genes that were especially affected by damaging genetic mutations, and grouped the disease genes according to their cellular functions. I contributed to the study by using the information about disease-causing mutations discovered in patients to pinpoint groups of early heart cell that are mostly affected by them, and found that the precursors of the heart muscle are the cellular hotspot of a specific type of congenital heart disease.

In the third study presented in this thesis, we used the same single-cell approach as in the first study, but here we identified a specific early heart cell type that can be found both in the human embryonic heart and in the *in vitro* heart development model system. This cell population is closely involved in building the complex structures of the heart, which connects the heart muscle to the great vessels, or in other words aorta and pulmonary arteries. As this structure is especially sensitive towards genetic mutations, our study identified a cell population with high relevance for modeling congenital heart defects.

To conclude, improving treatments and diagnostics for patients with complex congenital heart defects requires deep understanding of the normal heart development and genetics causing these defects on molecular and cellular level. In this doctoral thesis work, me and my colleagues built a roadmap of early heart cell differentiation (paper I); decoded the genetic mutations causing complex congenital heart defects and studied how they affect heart cells in early heart development (paper II); and identified a new cell population in the developing human heart, which may be closely involved in the early onset of congenital heart defects (paper III).

ABSTRACT

Cardiac progenitors are the fundamental building blocks of the heart and play a central role in the pathogenesis of congenital heart defects. Recent single-cell RNA sequencing studies have revealed that the cell populations in both developing and mature heart are more diverse than previously recognized. In particular, there has been a great interest in characterizing the transcriptional signatures of human cardiac progenitor populations and building a roadmap of the early heart lineages.

This thesis work is a part of an ongoing effort to map the early human heart progenitors, their developmental dynamics and their role in the pathogenesis of congenital heart defects at a single-cell resolution.

In **paper I**, we present a single-cell characterization of *in vitro* cardiac differentiation of human embryonic stem cells to beating cardiomyocytes and other populations. We used single-cell RNA sequencing data integration to compare the *in vitro*-derived cardiac cells to human embryonic heart, studied the developmental dynamics of cardiac progenitors by building a differentiation roadmap, and investigated the effect of loss of ISL1 transcription factor on the differentiation process.

In **paper II**, we mapped the genetic landscape of non-syndromic Tetralogy of Fallot, a form of complex congenital heart defect, in a cohort of 146 patient-parent trios. We intersected the identified disease-associated genes with single-cell RNA sequencing analysis of *in vivo* and *in vitro* human cardiac development generated by us and other laboratories. Interestingly, we could pinpoint cardiac progenitors as a cellular hotspot in cardiac disease pathogenesis.

Paper III introduces a novel cardiac progenitor marked by LGR5, which exists both in the human embryonic heart *in vivo* and the human embryonic stem cell cardiac differentiation system *in vitro*. The LGR5+ cono-ventricular progenitor population originates from the ISL1+ progenitor pool and populates the cardiac outflow tract, a cardiac region often affected in congenital heart defects.

Taken together, these studies support the importance of cardiac progenitors in the pathogenesis of cardiac abnormalities and explore their developmental dynamics at a single-cell level.

LIST OF SCIENTIFIC PAPERS

I. **Mononen M**[+], Leung CY[+], Xu J, Chien KR[*]. Trajectory mapping of human embryonic stem cell cardiogenesis reveals lineage branch points and an ISL1 progenitor-derived cardiac fibroblast lineage. Stem Cells. 2020 Oct 1;38(10):1267-1278.

II. Tang C[+], **Mononen M**[+], Hong H[+], Jin SC, Zhuang X, Lam W-Y, Garcia-Barceló M-M, Yang Y, Chien K, Tam P, Gruber P[*]. Whole genome sequencing of a Chinese cohort reveals insights into the genetic landscape and molecular mechanisms of Tetralogy of Fallot. *Manuscript*.

III. Sahara M[+*], Santoro F[+], Sohlmér J, Zhou C, Witman N, Leung CY, **Mononen M**, Bylund K, Gruber P, Chien KR[*]. Population and Single-Cell Analysis of Human Cardiogenesis Reveals Unique LGR5 Ventricular Progenitors in Embryonic Outflow Tract. Dev Cell. 2019 Feb 25;48(4):475-490.e7.

Scientific paper not included in this thesis:

Pölönen P, Jawahar Deen A[+], Leinonen HM[+], Jyrkkänen H-K, Kuosmanen S, **Mononen M**, Jain A, Tuomainen T, Pasonen-Seppänen, Hartikainen JM, Mannermaa A, Nykter M, Tavi P, Johansen T, Heinäniemi M, Levonen A-L[*]. Nrf2 and SQSTM1/p62 jointly contribute to mesenchymal transition and invasion in glioblastoma. Oncogene. 2019 Dec;38(50):7473-7490.

[+] These authors contributed equally

[*] Corresponding author

CONTENTS

LIST OF ABBREVIATIONS

Cas9	CRISPR associated protein 9
cDNA	Complementary DNA
CHD	Congenital heart defect
CRISPR	Clustered regularly interspaced short palindromic repeats
DORV	Double outlet right ventricle
DPC	Days post conception
ERCC	External RNA Controls Consortium
FACS	Fluorescence-activated cell sorting
FHF	First heart field
GATK	Genome Analysis Toolkit
HDR	Homology-directed recombination
hESC	Human embryonic stem cell
hiPSC	Human induced pluripotent stem cell
hPSC	Human pluripotent stem cell
HR	Homologous recombination
Indel	Insertion and deletion
ISL1	ISL LIM Homeobox 1 transcription factor
IWP	Inhibitor of Wnt ligand production
LGR5	Leucine rich repeat containing G protein-coupled receptor 5
NHEJ	Non-homologous end joining
PAM	Protospacer adjacent motif
PCGC	Pediatric Cardiac Genomics Consortium
PTA	Persistent truncus arteriosus
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
SHF	Second heart field

TGA	Transposition of great arteries
t-SNE	t-distributed stochastic neighbor embedding
UMAP	Uniform manifold approximation and projection
VEGF	Vascular endothelial growth factor
WNT	Wingless-type
WPC	Weeks post conception

1 INTRODUCTION

From the first heart beat to the last one, the human heart contracts more than two billion times over a lifetime. To get an early start, the heart is the first organ in an embryo to develop and initiate its function (Gilbert 2010). Building a heart is a complex process that depends on the differentiation and assembly of a family of cardiac lineages constituting its muscle compartments, vasculature, connective tissues and conductive system. These lineages appear early in embryonic development, pass through a sequence of progenitor stages and finally form the cell types found in the fully developed heart. (Meilhac and Buckingham 2018)

The cellular mechanisms steering heart progenitors through their development, including formation, proliferation, migration, differentiation, commitment and assembly, are regulated by specific signaling pathways, transcription factors and other mediators. Identification of the various cardiac progenitor populations and studying the molecular mechanisms regulating them is essential for understanding the developmental biology of the heart as well as the early pathogenesis of developmental cardiac malformations (Witman et al. 2020).

Congenital heart defect (CHD) is diagnosed in about 1% of all newborns and is therefore a significant health concern (Van Der Linde et al. 2011). Tetralogy of Fallot is the most common type of cyanotic CHD, and is characterized by the misalignment of the cardiac outflow tract, a complex developmental structure which later connects the cardiac ventricles and the great vessels (Apitz et al. 2009). Origins of complex CHD including Tetralogy of Fallot are not completely understood, but large cohort studies indicate that the cause is often genetic (Lalani and Belmont 2014).

A large body of research has been dedicated to understanding the disease mechanisms behind CHD phenotypes. These studies include genetic analysis of patients (Richter et al. 2020; Page et al. 2019; Jin et al. 2017; Lalani and Belmont 2014) and mechanistic studies on cells, tissues and animal models to pinpoint the connection between genetic mutations and altered cellular functions (Sharma 2020; Moon 2008). It is increasingly clear that results from animal models do not perfectly translate to human (Brawand et al. 2011; Lin et al. 2014; King and Wilson 1975). Therefore, there has been a large interest in generating cardiac progenitors and mature cells from human pluripotent stem cells (hPSCs) to be able to work on a human-based model system and access the developmental stages of cardiogenesis which cannot be accessed in human embryos (Protze et al. 2019). The hPSC-derived cardiac cells follow closely the developmental patterns of their counterparts in human embryonic heart (Später et al. 2014; Mummery et al. 2012), and therefore the hPSC cardiac differentiation provides a valuable model system for studying both cardiogenesis and CHD.

Single-cell transcriptomics is a state-of-the-art method for analysis of cells and tissues in health and disease together with other single-cell approaches. With single-cell RNA sequencing it is now possible to build a detailed and comprehensive roadmap of the cardiac lineages derived from hPSCs and human prenatal tissues (Paik et al. 2020).

Relying on single-cell RNA sequencing, we focused on building a connected roadmap of hPSC-based cardiac differentiation (paper I), mapping the genetic and cellular disease mechanisms of Tetralogy of Fallot by cross-referencing genomic sequencing and single-cell transcriptomics (paper II) and studying the cellular heterogeneity of hPSC-derived and human embryonic/fetal cardiac cells (paper III).

2 PROGENITORS AND EARLY HEART DEVELOPMENT

2.1 ANALYSIS OF HUMAN CARDIAC PROGENITORS

Human cardiac progenitors are the focus point of this thesis, yet they are difficult to access. A growing number of research groups are studying human pre-implantation embryos to answer some of the fundamental questions of the early human development (Xiang et al. 2020; West et al. 2019; Zhou et al. 2019; Stirparo et al. 2018; Boroviak et al. 2018; Petropoulos et al. 2016; Blakeley et al. 2015; Niakan and Eggan 2013; Yan et al. 2013). Culturing human embryos beyond the second week of embryonic development is ethically troublesome (Reyes and Lanner 2017), and consequently, human gastrulation and the subsequent events remain rather inaccessible for researchers. This is problematic for the study of human cardiogenesis, as the human heart development starts about 18 days after fertilization when the cardiac crescent arises (Sylva et al. 2013). Even studying the developmental stages which can be accessed in human (from about 4.5 weeks post conception [WPC]) is a challenge, as the human embryonic and fetal tissues available for researchers are sparse.

The discovery of hPSCs and understanding the fundamentals of the early heart development, discussed in the next sections, have allowed the derivation of human cardiac progenitors and cardiomyocytes using various differentiation protocols (Protze et al. 2019; Später et al. 2014; Mummery et al. 2012). This has greatly advanced our knowledge about the molecular mechanisms underlying the human-specific cardiac specification and disease development (Protze et al. 2019). Furthermore, the hPSC-based cardiac differentiation provides a platform for drug screening and development of cell-based therapies for cardiac regeneration (Protze et al. 2019; Parrotta et al. 2019; Ye et al. 2018; Sala et al. 2016; Drawnel et al. 2014) (Figure 1).

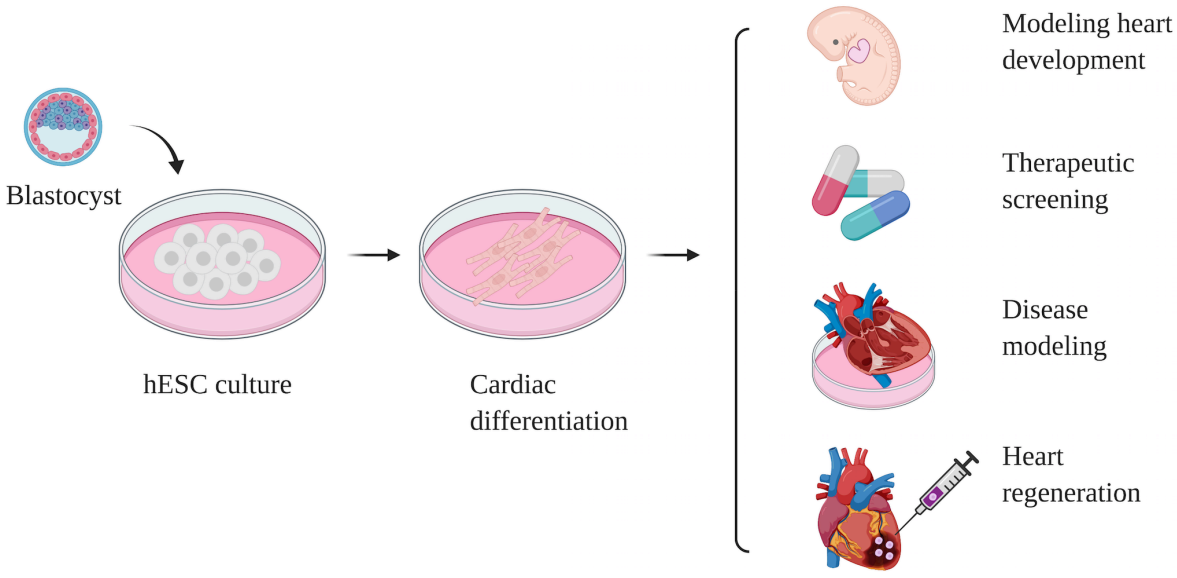

Figure 1. Derivation of hESCs from human blastocysts and the applications of hESC-derived cardiac cells. Created with BioRender.com.

With an access to a range of developmental stages of human cardiac progenitors, it is now possible to build a differentiation roadmap of early human heart development using single-cell RNA sequencing (Paik et al. 2020). In particular, single-cell RNA sequencing has proven useful for studying the emergence and segregation of cardiac progenitors and lineages, and the molecular mechanisms regulating these events (Protze et al. 2019). Moreover, in combination with CRISPR/Cas9-mediated gene editing, molecular mediators of human cardiac differentiation can be characterized and defined (Lian et al. 2014), which is particularly useful for understanding genetic mutations underlying CHD pathogenesis.

2.2 FROM MESODERM TO CARDIAC PROGENITORS

To be able to generate human cardiac progenitors and cardiac cell types from hPSCs, it is essential to understand how the cardiac lineages develop in an embryo (Protze et al. 2019). We know the anatomical timing of cardiogenesis in human embryos (Sylva et al. 2013), and animal models have helped filling in many of the missing details of cellular and molecular mechanisms taking place *in vivo*. Here, I will give an overview of the *in vivo* differentiation of cardiac cell populations and cardiac morphogenesis. The human developmental time points and cellular and molecular insights from animal studies run in parallel in this overview.

Human heart development starts from gastrulation at the end of the second WPC. Many biology students have heard the phrase "It is not birth, marriage, or death, but gastrulation, which is truly the most important time in your life" by Lewis Wolpert. During gastrulation, migration of cells through the primitive streak results in reorganization of a bilaminar embryonic disc, which makes the embryo within the blastula into a trilaminar disc composed of three germ layers: mesoderm, endoderm and ectoderm. Like most of the tissues of the cardiovascular system the heart mostly derives from lateral plate mesoderm, although derivatives of all three germ layers contribute to heart morphogenesis. (Meilhac and Buckingham 2018; Gilbert 2010) During gastrulation and the following events, Nodal signaling induces the expression of posterior-specific genes in the epiblast, which promotes the formation of mesoderm (Brennan et al. 2001). Canonical wingless type (Wnt)/β-catenin pathway induces the formation of *Brachuyry*/*T*-expressing mesoderm, and subsequent inhibition of canonical Wnt/β-catenin pathway and induction of non-canonical Wnt pathway leads to *Mesp1*-expressing cardiac mesoderm specification (Gessert and Kuhl 2010). Studies mapping early cardiac precursors in mouse and chicken have located the cardiogenic progenitors in the anterior part of the primitive streak (Garcia-Martinez et al. 1993; Tam et al. 1997; Kinder et al. 1999). Based on *Mesp1*-Cre-driven lineage tracing, approximately 70% of the heart is generated by the *Mesp1*-expressing mesoderm (Lescroart et al. 2014; Saga et al. 1999), including ventricular, atrial and outflow tract myocardium, endocardium, pericardium and the cardiac conduction system (Lescroart et al. 2018; Meilhac and Buckingham 2018).

Around third week of the human development, 17-19 DPCs, the first region showing cardiac specification and commitment, the cardiac crescent, forms as two symmetrical parts of the mesoderm adjacent to the underlying endoderm (Sylva et al. 2013). The differentiation of cardiomyocytes from the cardiac crescent progenitors depends on paracrine signaling from

the endoderm and ectoderm including bone morphogenetic protein (Bmp), sonic hedgehog (Shh), fibroblast growth factor (Fgf), Wnt and Notch families of growth factors (Witman et al. 2020; Harvey, 2002, Lough and Suki 2000; Zaffran and Frasch 2002). Furthermore, the foregut endoderm provides signals steering the direction of cardiac progenitor migration, and plays a major role in the formation and shape of the cardiac crescent (Schultheiss et al. 1995; Nascone and Mercola 1995; Lough and Sugi 2000). At the cardiac crescent stage, the progenitors express various levels of several cardiogenic transcription factor- (e.g. *Isl1*, *Nkx2-5*, *Mef2c* and *Gata4*), surface markers- (e.g. *Pdgfra* and *Flk1*) and early sarcomeric protein-coding genes (e.g. *Myl4*, *Myl7* and *Tnnt2*) (Behrens et al. 2013; Somi et al. 2004).

Folding of the embryo starts at the end of the third week, bringing the two lateral parts of the cardiac crescent together and forming the anterior part of the primitive heart tube (Sylva et al. 2013). Due to the automaticity of the early cardiomyocytes, the human heart tube starts to beat around 21-23 DPC (Vincent and Buckingham 2010). The heart tube is flanked by primitive inlet and outlets, the precursors of the inflow tract and outflow tract, bringing blood in and out of the developing heart (Sylva et al. 2013).

2.3 HEART FIELD PROGENITORS

Clonal analyses based on α-cardiac actin expression in mouse embryos revealed that the heart is mostly derived from two waves of cardiac progenitors (Meilhac et al. 2004). Subsequent studies have shown that those two populations are spatiotemporally guided by chemotaxis and regulated by cardiac transcription factors (Xiong et al. 2019; Waardenberg et al. 2014). The two cardiac progenitor lineages, first one giving rise to the linear heart tube and the second one gradually migrating to the poles of the heart tube, are called first heart field (FHF) and second heart field (SHF), respectively (Moorman et al. 2007; Buckingham et al. 2005; Kelly et al. 2001; Mjaatvedt et al. 2001; Waldo et al. 2001). The primitive heart tube generates the left ventricle and contributes to other parts of the heart, whereas cells migrating to the heart tube from the SHF generate the major part of the atria, outflow tract and right ventricle (Zaffran et al. 2004; Cai et al. 2003). The two cardiac lineages segregate as early as at the onset of gastrulation or even earlier (Meilhac et al. 2004), and show different developmental dynamics; the FHF progenitors directly commit to the cardiomyocyte lineage, while the SHF progenitors undergo a sequence of gradual commitment and differentiation (Xiong et al. 2019). Yet, it is still not clear whether the FHF and SHF differ in their fundamental progenitor potential or whether the differences between them are only due to their location in the developing heart (Tyser et al. 2021).

Many of the early cardiac lineage marker genes, such as *Gata4* and *Nkx2-5*, are expressed in both heart fields (Meilhac and Buckingham 2018), and definitive markers for the FHF are somewhat lacking (Buckingham et al. 2005). More specific to the FHF derivatives, transcription factor *Hand1* has been shown to be expressed and regulate the left side of the developing murine heart (Biben and Harvey 1997; McFadden et al. 2005). Similarly, *Tbx5* mutant mice display hypoplasia of the left ventricle and atria, indicating an important role in the FHF derivatives (Bruneau et al. 2001). However, neither *Hand1* or *Tbx5* can be

considered exclusively as FHF markers, as both of them are expressed by subsets of SHF-derivatives (Barnes et al. 2010; Rana et al. 2014).

In contrary to the FHF, the SHF progenitors can be distinguished by their expression of specific genes, such as *Isl1*, *Fgf8*, *Fgf10* and *Tbx1* (Theveniau-Ruissy et al. 2008; Kelly et al. 2001; Buckingham et al. 2005; Dyer and Kirby 2009; Watanabe et al. 2010). At the progenitor stage, they compose a rapidly proliferating and migratory population which is added to the primitive heart tube, whereas the differentiated myocytes do not contribute to the heart growth by division to any significant extent (van den Berg et al. 2009; Kelly et al. 2001; Mjaatvedt et al. 2001; Waldo et al. 2001; Cai et al. 2003). These developmental phases coincide with the expression of *Isl1* and *Nkx2-5*, respectively (Buijtendijk et al. 2020). In particular, Cre-mediated lineage tracing in mouse revealed, that the transcription factor ISL LIM Homeobox 1 transcription factor (ISL1) broadly marks the SHF progenitors (Cai et al. 2003), although another study showed that some of the cardiac Isl1-derivatives are of neural crest origin (Engleka et al. 2012). Lineage tracing and clonal assays of mouse cardiac progenitors have further shown that the *Isl1*-expressing SHF-derivatives include myocardial as well as endothelial and smooth muscle cells (Moretti et al. 2006; Laugwitz et al. 2005). This is also true for ISL1-cardiac progenitors derived from hPSCs (Bu et al. 2009). The hPSC-derived ISL1+ cardiac progenitors are of specific interest for regenerative medicine applications, as they have been shown to generate ventricular-like progenitors, with potential to engraft, expand, differentiate, self-assemble and maturate upon transplantation into mouse heart and kidney (Foo et al. 2018).

It should be noted that although the ISL1+ heart progenitors have been identified in human prenatal heart (Yang et al. 2013; Bu et al. 2009), it is not yet completely clear what the species differences are and if the human and mouse heart fields behave in a similar manner (Pandur et al. 2013).

2.4 OTHER PROGENITORS AND LINEAGES

The majority of the cells of the heart are cardiomyocytes (Zhou and Pu 2016). However, connective tissue cells, namely fibroblasts, smooth muscle cells and endothelial cells, comprise a significant portion of the volume of the heart and carry out various important functions (Zhou and Pu 2016; Banerjee et al. 2007). In addition to the heart field progenitors, other lineages are required to build these cell types, including the neural crest. Endocardium, pericardium and myocardium arise from the cardiogenic *Mesp1*-expressing mesoderm and share common markers, such as *Isl1* and *Nkx2-5* (Ma et al. 2008; Sun et al. 2006; Cai et al. 2003; Saga et al. 2000; Saga et al. 1999; Mikawa et al. 1992; Kaufman and Navaratnam 1981), while cardiac neural crest cells are derivatives of the neuroectoderm (Jiang et al. 2000).

Endocardial cells first generate the inner epithelial lining of the myocardial heart tube, and later the inner layer of the heart chambers (Sylva et al. 2013). Both myocardial and endocardial progenitors express *Isl1* and *Nkx2-5* (Ma et al. 2008; Cai et al. 2003), whereas

Nfatc1 specifically marks the endocardial lineage (Wu et al. 2011; de la Pompa et al. 1998). Between the myocardial and endocardial components lies cardiac jelly, which is a layer of extracellular matrix (Little et al. 1989). Signals from the surrounding myocardium, including Bmp and transforming growth factor beta (Tgfβ), induce epithelial-to-mesenchymal transition within the endocardial population, which leads to the formation of endocardial cushions (Azhar et al. 2003; Kruithof et al. 2012; Brown et al. 1996; Eisenberg and Markwald 1995; Potts and Runyan 1989). The endocardial cushion further forms the membranous parts of the outflow tract, atrial and ventricular septa, the tricuspid valve and the mitral valve (Kruithof et al. 2012).

During heart looping around 22 to 23 DPC, the epicardial cells start to populate the outer surface of myocardium generating the epicardium (Sylva et al. 2013). Epicardial cells originate from the pro-epicardial organ located on the inflow tract side of the heart, and the epicardial precursors are specifically marked by *Tbx18* and *Wt1* expression (Meilhac and Buckingham 2018). The origin of the pro-epicardial organ cells has been unclear, but a recent single-cell RNA sequencing study identified a novel *Mesp1*-mesoderm-derived progenitor population (juxta-cardiac field), which gives rise to both cardiomyocytes and epicardial cells (Tyser et al. 2021). In addition to covering the heart, epicardial cells give rise to cardiac interstitial fibroblasts, smooth muscle and endothelial cells of coronary arteries, valvular interstitial cells and annulus fibrosus, i.e. the fibrous tissue insulating atrial and ventricular electrical currents (Dettman et al. 1998; Mikawa and Gourdie 1996; Perez-Pomares et al. 1997; Poelmann et al. 1993). Most of the cardiac fibroblasts, especially the cardiac interstitial fibroblasts, originate from a subpopulation of the epicardial cells undergoing epithelial-to-mesenchymal transition, although some are generated by the cardiac neural crest and endocardium (Meilhac and Buckingham 2018; Ivey and Tallquist 2016; Ali et al. 2014).

Cardiac neural crest is a multipotent migratory progenitor population marked by *Wnt1*-expression, which differentiates into smooth muscle at the base of the large vessels, connective tissue and neurons of outflow tract, fibroblasts of the right atrium, aorticopulmonary ridge and cardiac parasympathetic nervous system, as well as contributes to the cardiac septa and mesenchyme of the aortic and pulmonary valves (Kirby et al. 1983; Bergwerff et al. 1998; Creazzo et al. 1998; Jiang et al. 2000; Jain et al. 2011). Retrovirus labelling has indicated, that neural crest cells gradually migrate to the heart region, first populating the outflow tract septum and pharyngeal arch arteries, and later contribute to the remodeling of the proximal pharyngeal arch arteries (Boot et al. 2003). Contributions of the neural crest to heart and large vessel development can be summed up in three main points based on neural crest ablation studies: 1) it participates in building the septum of the cardiac outflow tract (Kirby et al. 1983), 2) modulates FGF signaling, which affects both pharyngeal and heart development (Farrell et al. 2001), and 3) is indirectly involved in the recruitment of SHF contributing to outflow tract elongation (Yelbuz et al. 2002; Waldo et al. 2005).

2.5 CARDIAC PATTERNING AND MORPHOGENESIS

The above mentioned lineages contribute to the complex events following the formation of a simple heart tube, and subsequent cardiac patterning and morphogenesis. Regional gene expression, as well as the hemodynamic forces generated by the beating heart tube and blood circulation contribute to the development of cardiac chambers, valves and conduction system. The specific properties of cardiomyocytes in the different heart compartments are established by differential gene expression of transcription factors, signaling molecules, structural proteins and ion channels (Sylva et al. 2013; Christoeffels et al. 2000; Bruneau 2013). In particular, a group of T-box transcription factors contribute to cardiomyocyte fate determination and patterning of cardiomyocytes with different electrophysiological properties (Sylva et al. 2013), whereas the anterior-posterior patterning, i.e. atrial and ventricular separation, is established by retinoic acid-induced Hox genes (Bertrand et al. 2011).

Corresponding to the fourth week of human development, the heart tube elongates as a result of SHF migration, and starts to bend into an S-shape. Simultaneously with the elongation, rightward looping of the heart tube begins, as does ballooning and development of the future ventricles (Sylva et al. 2013). Soon after ventricular and atrial ballooning, septation starts to separate the four cardiac chambers. The cardiac septa are composed of both muscular and membranous parts (Franco et al. 2006; McFadden et al. 2005; Habets et al. 2002), and defects in septation are common representations of CHD, often affecting the membranous part of cardiac septa (Murray 1948). Towards the end of the sixth WPC, the ventricles, atria and outflow tract of the human heart are separate structures located in their final positions, but will continue to grow and mature (Sylva et al. 2013).

The cardiac outflow tract develops from a myocardial tube into separate pulmonary and aortic outlets from 25 to 48 DPC in the human developmental timescale (Sylva et al. 2013). At the early stages of heart development, the outflow tract wall is exclusively composed of myocardium, whereas the lumen contains the cardiac jelly and endocardium which will give rise to the outflow tract cushions (Anderson et al. 2016). The outflow tract elongates as a result of SHF migration, and as the development proceeds, SHF continues to add both myocardial and non-myocardial components to the outflow tract (Kelly et al. 2001; Mjaatvedt et al. 2001; Waldo et al. 2001; Cai et al. 2003; Rana et al. 2014). Migration of cells from the neural crest into the heart is crucial for the outflow tract development, and is regulated by Wnt, Bmp, Fgf and Semaphorin signaling (Stoller and Epstein 2005). Within the outflow tract lumen, neural crest derivatives and the endocardial cushions build a septum as well as the semilunar valves (De Lange et al. 2004). The septum undergoes a 180˚ counterclockwise rotation in a spiral-like motion, resulting in the formation of the two large cardiac vessels. In particular, the SHF adds myocardial cells to the spiraling outflow tract, which is required for the appropriate alignment of the aorta and pulmonary trunk relative to the left and right ventricles (Ward et al. 2005). Following the fusion of the outflow tract cushions, the aorta and pulmonary trunk get separated gradually in distal to proximal order. Furthermore, remodeling of three pairs of pharyngeal arch arteries (3rd, 4th and 6th) contributes to the final morphology of the aortic arch and pulmonary trunk. (Anderson et al. 2016; Sylva et al. 2013)

In addition to Tgfβ and Bmp which regulate the epithelial-to-mesenchymal transition of cardiac cushions, Vegf and its downstream Notch signaling pathways play an important role in the outflow tract septum development, by regulating both endocardium and the SHF myocardium (Dor et al. 2001; van den Akker et al. 2012; Niessen and Karsan 2008).

3 CONGENITAL HEART DEFECTS

3.1 INTRODUCTION TO CONGENITAL HEART DEFECTS

Cardiogenesis is established by a series of events involving integration of multiple cell types, and dysregulation of these processes results in morphological defects. CHD is the most common type of birth anomaly, affecting approximately 1% of newborns (Shabana et al. 2020; Liu et al. 2019; van der Linde et al. 2011, Hoffman et al. 2002), and has the estimated prevalence of 2-3% (Pierpont et al. 2018). When including stillbirths that are caused by severe CHD, the percentage of CHD cases is significantly higher (Marelli et al. 2014; Hoffman 1995). The reported total CHD prevalence, especially that of mild congenital cardiac lesions, has increased over time due to improved clinical diagnostics and treatment options. As a consequence of increased survival beyond childhood, there are more adults than children with CHD, which has important implications for the modern societies (Liu et al. 2019; Marelli et al. 2014, van der Linde et al. 2011). The manifestations of CHD include a range of cardiac malformations, the most common types globally being ventricular and atrial septal defects, patent ductus arteriosus, pulmonary stenosis, Tetralogy of Fallot and transposition of the great arteries (Liu et al. 2019; van der Linde et al. 2011). Because of its frequency and impact on the life of the patients, CHD is an important health concern and branch of cardiovascular research.

Exposure to environmental factors, such as maternal retinoic acid intake, diabetes, smoking and exposure to medications and chemicals are implicated in both CHD and other common developmental defects, such as spina bifida and cleft lip and palate (Zhu et al. 2009; Hinton 2013; Jenkins et al. 2007). However, cardiogenesis is particularly sensitive to genetic disturbances, and many epidemiological studies have established that genetics is the main factor behind CHD pathogenesis (Shabana et al. 2020). Large epidemiological studies indicate, that a single genetic cause is identified in only 20-30% of CHD patients, whereas a large portion of the cases is caused by rare variants and combinations of damaging mutations (Pierpont et al. 2018, Cowan and Ware 2015). In other words, CHD is characterized by a highly complex and contextual genetic landscape.

3.2 GENETIC MECHANISMS

Approximately 400 genes have been attributed to CHD (Williams et al. 2019). The disease-causing variants affect a heterogeneous group of genes, in particular those coding for transcription factors, signaling and adhesion molecules, structural proteins, chromatin modifiers and cilia (Nees and Chung 2020), which can interfere with normal progenitor specification, differentiation, migration and patterning. Although a large number of gene variants have been shown to cause heart anomalies, the mechanisms that underlie the genotype-phenotype connection remain unclear for many types of complex CHD.

A specific case of CHD can be 1) monogenic or polygenic, meaning that the phenotype is caused by only one genetic mutation or a combination of several genetic mutations, 2) syndromic or non-syndromic, meaning that the cardiac phenotype can appear in combination

with extracardiac anomalies or alone, as well as 3) familial or sporadic, meaning that the susceptibility gene mutation and cardiac phenotype run in the family following Mendelian heritage, or the patient phenotype is caused by a rare recessive or *de novo* mutation in a proband whereas the parents are unaffected (Fahed et al. 2013). Majority of the CHD cases, about 80%, are non-syndromic (Nees and Chung 2020). Furthermore, CHD cases are predominantly sporadic, and major part of the sporadic CHD cases is estimated to be explained by *de novo* damaging mutations (Shabana et al. 2020).

Genetic variants can be divided into structural variants and point mutations (Shabana et al. 2020). Structural variants include chromosomal aneuploidies and copy number variants. Trisomy 13, 18 and 21 are examples of chromosomal aneuploidies which are characterized by a range of phenotypes affecting the heart and other organ systems (Calcagni et al. 2017). The abnormal number of chromosomes in these patients causes changes in the expression dosage of thousands of genes (Wimalasundera and Gardiner 2004). The literature gives multiple definitions for copy number variants, but in general it can be said that they are sequences of deletions or multiplications. Due to often affecting the expression of multiple genes, copy number variants are commonly associated with syndromic CHD, but can also cause non-syndromic forms (Nees and Chung 2020).

Point mutations can be either single nucleotide polymorphisms or insertions and deletions (indels) of short sequences, and cause both syndromic or non-syndromic forms of heart disease. Moreover, both coding and non-coding point mutations can cause CHD (Richter et al. 2020). When in the coding region, small mutations can cause a frameshift leading to a premature stop codon or disrupt important functional sites. Single nucleotide polymorphisms can also be non-synonymous, changing the amino acid sequence or synonymous which might affect the post-transcriptional processing of the mRNA molecule. The role of damaging variants within the non-coding regions in CHD is not yet well understood, but mutations in splice junctions and splice enhancers, sequences regulating RNA homeostasis and localization, regulatory RNAs, promoters, and enhancers can lead to dysregulation of gene expression. (Chahal et al. 2019, Ward and Kellis 2012)

The perturbations in heart development by damaging genetic mutations is often caused by haploinsufficiency, i.e. the mutation causes either loss of gene expression, decreased expression or decreased protein function (Fahed et al. 2013). Less frequently, a mutation can lead to increased gene activity, which nevertheless can be equally damaging for cellular functions as loss of function mutations (Shabana et al. 2020).

Disease-causing gene mutations sometimes interact with other stochastic genetic and environmental factors modifying the patient phenotype, although the molecular mechanisms behind these interactions are not well understood (Akhirome et al. 2017; Gifford et al. 2019). For example, the length of the DiGeorge associated 22q11.2 microdeletion does not directly correlate with the patient phenotype, which, instead, is likely determined by genetic modifiers in other genomic loci (Aggarwal et al. 2008). Furthermore, damaging mutations in several different genes can produce the same cardiac phenotype, and the same mutation in two

individuals may not cause the same heart defect (Shabana et al. 2020). It will be the challenge for the coming decades to understand how molecular pathways and gene families interact with different modifiers and genetic backgrounds to result in specific patient phenotypes (Jenkins et al. 2007, Zhu et al. 2009).

3.3 RECENT GENOME-WIDE CARDIAC COHORT STUDIES

Next-generation sequencing-based genetic studies have rapidly identified numerous novel disease-associated genes and gene variants. The next-generation sequencing approaches include whole-exome sequencing and whole-genome sequencing, which cover genetic variations within the coding genome or the complete genome, respectively.

A recent large whole-exome sequencing study by the Pediatric Cardiac Genomic Consortium (PCGC) mapped novel damaging rare and dominant inherited and *de novo* mutations in a cohort of 2,871 CHD probands, including 2,645 family trios (Jin et al. 2017), being the most comprehensive published genetic study on a single CHD cohort to date. Among other findings, the study reported a recessive founder mutation in *GDF1* accounting for ~5% of severe CHD in Ashkenazim population, recessive *MYH6* variants accounting for ~11% of Shone complex cases, and dominant mutations in *FLT4* accounting for ~2% of Tetralogy of Fallot (Jin et al. 2017). Despite the large number of patients included in the study, it only covers a fraction of the total genetic causes of CHD. Sifrim et al. estimated that whole-exome sequencing of ~10,000 family trios would be needed to cover most of the dominant CHD-associated genes (Sifrim et al. 2016, Jin et al. 2017).

The first large whole-genome sequencing trio study included 749 CHD probands with parents, and used neural network prediction and enhancer analysis to identify potentially damaging mutations within the non-coding genome (Richter et al. 2020). The study identified non-coding variants within human fetal cardiac regulatory regions based on H3K27ac chromatin immunoprecipitation of human fetal cardiac tissue, and functionally validated selected mutated regulatory sequences by massively parallel reporter assay, a high-throughput platform for measuring regulatory activity of thousands of normal and mutated regulatory DNA sequences in parallel in an *in vitro* system (Melnikov et al. 2014). The study established that non-coding *de novo* variants significantly contribute to CHD pathogenesis, but points out that the genetic mechanisms remain to be explored. Several high-throughput technologies can be used to navigate the often cell type-specific functional implications of non-coding variants, including ChIP-seq, RNA-seq, ATAC-seq, DNAse-seq, MNase-seq and FAIRE-seq (Chahal et al. 2019).

3.4 MALFORMATIONS OF THE CARDIAC OUTFLOW TRACT

Defects in the outflow tract development represent approximately 30% of all CHD cases (Thom et al. 2006). Since the outflow tract development requires precise coordination of migratory progenitor populations of the SHF, endocardium and cardiac neural crest to establish elongation, septation, rotation and separation of the large vessels, genetic mutations impairing the function of any of these lineages may lead to disturbances in the morphogenetic

patterning of the outflow tract (Neeb et al. 2013). The major abnormalities of the outflow tract found in patients are:

Persistent truncus arteriosus (PTA): The aorta and the pulmonary trunk are not completely separated. Appears always together with atrioventricular septal defect.

Double outlet right ventricle (DORV): Both the aorta and pulmonary trunk are connected to the right ventricle as a result of partially failed outflow tract rotation. DORV always occurs with a ventricular septal defect connecting the ventricles.

Transposition of the great arteries (TGA): The aorta exits the right ventricle and the pulmonary trunk exits the left ventricle as a result of failed outflow tract rotation.

Tetralogy of Fallot: Misaligned outflow tract with associated overriding aorta, pulmonary stenosis, right ventricular hypertrophy and ventricular septal defect. (Neeb et al. 2013)

Other outflow tract -related anomalies include coarctation of the aorta, vascular ring, aberrant subclavian artery, pulmonary stenosis, patent ductus arteriosus, persistent left superior vena cava, anomalous pulmonary venous connection and various outflow tract valve deformities (Neeb et al. 2013). The disease mechanisms behind these malformations are not completely understood.

3.5 TETRALOGY OF FALLOT

3.5.1 Introduction to Tetralogy of Fallot

Tetralogy of Fallot represents 7-10% of all CHD cases (Ferencz et al. 1985; Bailliard and Anderson. 2009) and mainly affects the right ventricular outflow tract. The condition manifests as a combination of four cardiac structural abnormalities: pulmonary outflow tract obstruction (either infundibular narrowing or pulmonary valve stenosis), ventricular septal defect, overriding aorta, and right ventricular hypertrophy (Fallot. 1888) (Figure 2). Although described as four separate structural malformations, the phenotype has one cause, as the misaligned outflow tract is mainly responsible for the pulmonary outflow tract obstruction, ventricular septal defect and overriding aorta, and right ventricular hypertrophy is the consequence of the hemodynamic changes caused by these structural defects (Wiputra et al. 2018; Bailliard and Anderson 2009; Andreson and Weinberg. 2005; Van Praagh et al. 1970). From the cell biology perspective, the failure of the SHF progenitors to muscularize the outflow tract is a central issue in the Tetralogy of Fallot pathophysiology (Bailliard et al. 2009).

Cyanosis, caused by deoxygenated blood from the right ventricle entering the systemic circulation through aorta, is the clinical hallmark of Tetralogy of Fallot pathology. Consequently, the blood oxygen saturation is low, which can be observed as a bluish skin color of infants suffering from Tetralogy of Fallot (Figure 2). The Tetralogy of Fallot anatomy varies between patients, and the classical form of Tetralogy of Fallot is often seen together with other anomalies including pulmonary atresia, absent pulmonary valve, DORV,

atrioventricular septal defect, atrial septal defect and right aortic arch (Bailliard and Anderson. 2009). Tetralogy of Fallot patients usually undergo corrective surgery soon after birth, ideally within the first 12 months, and the choice of surgical intervention is determined by the severity of the cardiac phenotype (Bailliard and Anderson. 2009). Complete cardiac repair involves insertion of two patches; one to close the opening between ventricles and one to reconstruct the right ventricular outflow tract area (Kurosawa et al. 1988; Kurosawa et al. 1998). In the palliative surgical option, shunts are generated to connect systemic and pulmonary circulation (Sharkey and Sharma 2012). The corrective surgery can lead to complications, and even after successful intervention many patients suffer from persistent arrhythmias and dysfunction of the right ventricle due to backflow of blood from the pulmonary artery back into the right ventricle (Kurosawa et al. 1998). However, before the surgical correction was introduced in 1955, the mortality of untreated Tetralogy of Fallot patients was 50% in the few first years of life, which has dropped significantly with the development and improvement of surgical interventions (Apitz et al. 2009).

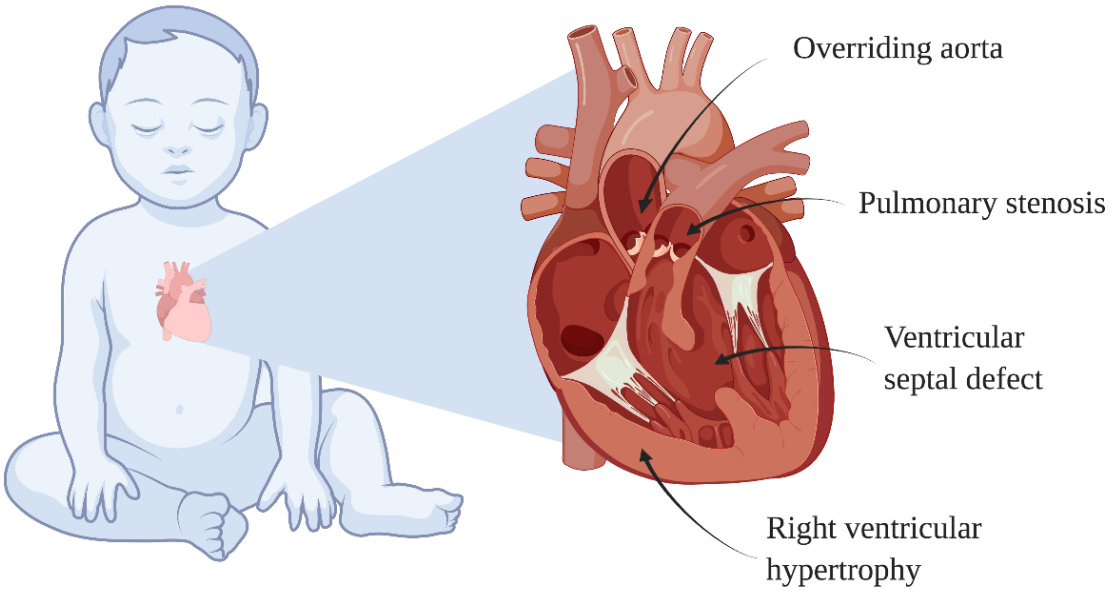

Figure 2. The major characteristics of Tetralogy of Fallot. Created with BioRender.com.

3.5.2 Genetics of Tetralogy of Fallot

The etiology of Tetralogy of Fallot is not completely understood, but in the light of present evidence, genetic predisposition plays a significant role in the disease pathogenesis (Morgenthau and Frishman 2018). The reason for uncertainty about the etiology of Tetralogy of Fallot and other complex CHD is that they are multifactorial, with both complicated genetics and involvement of environmental risk factors.

Several sequencing studies have identified genetic mutations affecting members of NOTCH and VEGF signaling pathways, particularly the VEGF receptors FLT4 and KDR, as well as NOTCH1, NOTCH2 and JAG1 from the NOTCH family, as primary candidates in the

pathogenesis of Tetralogy of Fallot (Reuter et al. 2019; Jin et al. 2017; Page et al. 2019). The VEGF pathway has been long known to play an important role in cardiovascular development (Carmeliet. 2000; Ferrera et a. 2003; Ylä-Herttuala et al. 2007). NOTCH signaling is closely intertwined with the VEGF pathway, being one of its main downstream pathways (Lawson et al. 2003; van den Akker et al. 2012). The importance of NOTCH and VEGF in Tetralogy of Fallot pathology is also supported by their eminent, but not completely understood, role in the development of the cardiac outflow tract (van der Akker et al. 2012).

In addition to signaling molecules, several transcription factors have been implicated in the pathogenesis of Tetralogy of Fallot. Much of the current knowledge about the role of certain transcription factors in Tetralogy of Fallot genetics comes from familial studies. For example, damaging mutations in *GATA4*, *TBX5* and *NKX2-5* have been found to result in familial isolated Tetralogy of Fallot in several instances (Dixit et al. 2018; Yang et al. 2013; Goldmuntz and Benson, 2001; Benson et al. 1999). The transcription factor-coding genes mentioned here are but a few examples, and many more have been implicated in the Tetralogy of Fallot phenotype, including *TBX1*, *FOXC1*, *FOXC2*, *FOXH1*, *HAND2*, *ZFPM2* and others (Morgentha and Frishman. 2018; Töpf et al. 2014; Zhang et al. 2014).

Approximately 20-30% of Tetralogy of Fallot cases are associated with genetic syndromes and rare diseases, including DiGeorge, Trisomy 21 (or Down syndrome), CHARGE, Opitz, Cantrell, Trisomy 22, Alagille, Goldenhar, Sotos, Torell, Noonan, Williams, Waarndenburg, Kabuki, Klippel Feil and cystic fibrosis (Morgenthau and Frishman 2018; Digilio et al. 1996). Despite the diversity of genetic syndromes associated with Tetralogy of Fallot, the genetics of syndromic cases is relatively better understood than that of the non-syndromic cases. Both mutations in the *TBX1* gene and its deletion or duplication in the 22q11.2 locus, which includes *TBX1* and other genes, in DiGeorge syndrome usually manifest with Tetralogy of Fallot (Merscher et al. 2001; Lindsay et al. 2001; Hasten et al. 2018). DiGeorge syndrome is strongly linked with the haploinsufficiency of *TBX1* and other genes within the associated locus, but mutations in other genetic loci have been found to modify the phenotype (Du et al. 2020). As another example, heterozygous mutations in *JAG1* and *NOTCH2* genes cause Alagille syndrome, and similar to DiGeorge, the phenotypic penetrance varies between patients (Mitchell et al. 2018; McDaniell et al. 2006). In the light of cohort studies, approximately 3-4% of Down syndrome patients suffer from Tetralogy of Fallot (Stoll et al. 2015; Källén et al. 1996).

The majority of Tetralogy of Fallot cases are isolated, and the genetic etiology of isolated Tetralogy of Fallot is complex and characterized by a plethora of rare genetic variants, which is true for isolated CHD in general (Nees and Chung. 2020). Discovery of genes causing non-syndromic Tetralogy of Fallot is largely facilitated by whole-exome and whole-genome sequencing studies based on large and clinically well-defined patient cohorts. The largest such study to date included whole exome sequencing of 829 non-syndromic Tetralogy of Fallot patients, and found that 7% of the study subjects harbored deleterious variants in *NOTCH1* and *FLT4* loci (Page et al. 2019).

3.5.3 Modeling the pathophysiology

Early embryonic manipulation studies, in particular using chick embryos as a model, revealed the importance of the SHF and neural crest in the outflow tract morphogenesis. These studies revealed that removal of the neural crest population before its migration to the outflow tract disrupts the formation of the aorticopulmonary septum leading to either common arterial outflow tract channel (i.e. PTA) or TGA (Kirby et al. 1983). Furthermore, neural crest plays another important indirect role in the outflow tract morphogenesis, as signaling by the neural crest is required for the recruitment of the SHF-derived myocardium into the developing outflow tract (Waldo et al. 2005). Consequently, the SHF fails to add cells to the outflow tract myocardium in chick embryos after cardiac neural crest ablation (Waldo et al. 2005). It was also shown that the direct disruption of the outflow tract-destined SHF population induces a variety of outflow tract malformations including TGA as well as reduced size of the pulmonary trunk, which is the hallmark of Tetralogy of Fallot pathogenesis (Ward et al. 2005).

Several genetically engineered mouse models have been developed to study specific cell lineages contributing to outflow tract development. For example, global deletion of *Tfap2a* (Brewer et al. 2002) and *Cited2* (Bamforth et al. 2001) affects the cardiac neural crest specification; deletion of *Pitx2* (Ai et al. 2006) and *Smad4* (Azhar et al. 2010) affects outflow tract myocardium; and deletion of *Nrp1* (Kawasaki et al. 1999) or *Sox4* (Schilham et al. 1996) affects outflow tract endothelium and endocardium. Moreover, various Cre recombinase mice have been generated for studies investigating the lineages contributing to outflow tract development. For instance, mice with Cre-knockin in the *Wnt1* (Jiang et al. 2000) and *Tfap2a* (Macatee et al. 2003) loci have been used to study the neural crest lineage; in *Isl1* (Park et al. 2006) and *Nkx2-5* (Lints et al. 1993) loci to study the myocardial lineage; and in *Tie1* (Gustafsson et al. 2001) and *Kdr* (Motoike et al. 2003) loci to study the endocardial lineage.

Multiple loss- and gain-of-function mouse models show the pathophysiology similar to human Tetralogy of Fallot (Snider and Conway 2011), and have revealed a few different explanations for the outflow tract phenotype. However, as human patients, genetically engineered mice usually show a spectrum of cardiac phenotypes despite sharing the same genotype. For example, *Hey2* knockout mice, which codes for a transcription factor modulating Notch signaling, display a variety of heart malformations including Tetralogy of Fallot (Donovan et al. 2002). Similarly, knockout and/or haploinsufficiency studies on *Nkx2-5*, *Tbx1* and *Jag1* have confirmed their role in outflow tract morphogenesis and Tetralogy of Fallot pathogenesis (Ashraf et al. 2014; Rana et al. 2014; Hofmann et al. 2012). In VEGF120/120 knockin mice (which only express the VEGF120 isoform), a localized increase of Vegf signaling in the SHF-derived outflow tract myocardium leads to perturbations in Notch signaling and outflow tract cushion hypoplasia. This mouse model has been shown to produce the Tetralogy of Fallot phenotype with high penetrance (Rammeloo et al. 2015; van den Akker et al. 2007). Furthermore, knockout of *Zfpm2* gene, a co-factor of Gata transcription factors, produces a Tetralogy of Fallot-like phenotype in mice, but also is

embryonic lethal and is accompanied by various extracardiac defects (Tevosian et al. 2000). These animal models have verified that the decreased outflow tract elongation causes the lack of counterclockwise rotation of the outflow tract cushions, which is the root of Tetralogy of Fallot pathogenesis (Snider and Conway 2011; Restivo et al. 2006). These examples further illustrate, that disturbances in various developmental mechanisms can cause the typical Tetralogy of Fallot phenotype, which applies to human Tetralogy of Fallot patients as well as mouse models.

In addition to animal models, cardiac differentiation of hPSCs has been increasingly used as a human-based alternative to study CHD (Matsa and Denning 2012; Ye et al. 2018; Shaheen et al. 2017). In particular, the development of the CRISPR/Cas9 system (Torres-Ruiz and Rodriguez-Perales 2017) has enabled efficient induction and correction of patient-specific mutations in hPSCs. In a recent study, Grunet and colleagues generated patient-specific induced pluripotent stem cells (iPSCs) from dermal fibroblasts collected from two patients with non-syndromic Tetralogy of Fallot (Grunet et al. 2020). Upon *in vitro* cardiac differentiation in a 2D culture system, the patient iPSC-derived cells showed significant decrease in the expression of cardiac transcription factors (e.g. *ISL1* and *HES1*), genes belonging to the NOTCH signaling pathway (e.g. *JAG1*), cell-cell adhesion molecules (e.g. *DSG2*) and several collagen-coding genes compared to the iPSCs derived from healthy controls. Another study showed, that pathogenic mutations of the NOTCH signaling-related gene *ADAM17* found in Tetralogy of Fallot patients induced hypertrophy of hESC-derived cardiomyocytes (Xie et al. 2019). The phenotype of *ADAM17* mutant cells was related to disturbances in heparin-binding EGF-growth factor signaling (Xie et al. 2019), which has been shown to regulate cardiac development and hypertrophy (Nanba and Higashiyama 2004).

4 SINGLE-CELL TRANSCRIPTOMICS AND HEART DEVELOPMENT

Single-cell transcriptomics provides an approach with great resolution and throughput for studying the patterns of gene expression that define the different stages of cardiac lineage development, and has facilitated the identification of novel and rare cell types in pre- and post-natal human hearts (Paik et al. 2020). To set the main stage of my thesis work, I will here give an overview of the current single-cell coverage of the pre-natal heart development with notes to the cell composition in mouse and in the adult human heart.

4.1 SINGLE-CELL ANALYSIS OF EARLY MOUSE HEART

The mouse model for early embryonic development is well established, and single-cell studies on mouse embryos have investigated the genetic regulation from implantation to gastrulation between embryonic day (E) 3.5 to 6.5 (Mohammed et al. 2017); conversion of pluripotent cells to *Mesp1*-expressing mesoderm around gastrulation at E6.25, 6.5 and 7.25 (Lescroart et al. 2018; Scialdone et al. 2016); lineage commitment of *Isl1*- and *Nkx2-5*-expressing populations through E7.5 to 9.5 (Jia et al. 2018); spatial specification of cardiac progenitors at E8.5, 9.5 and 10.5 (Li et al. 2016); and the anatomical maturation of primordial heart tube at E9.5 into a fully developed heart at post-natal day 21 (DeLaughter et al. 2016). Combining gene ablation with single-cell analysis, recent studies have elucidated the specific roles of *Hand2* in heart development and outflow tract defects covering E7.75 and 8.25 (de Soysa et al. 2019) and *Pitx2* in both heart fields at E10.5 and 13.5 (Kadow et al. 2019). Furthermore, single-cell analysis of the outflow tract cell types by Liu and colleagues, with an emphasis on smooth muscle cell development and trans-differentiation of myocardium to smooth muscle, provides a reference of the normal outflow tract development for studies on conotruncal heart defects (Liu et al. 2019). Interestingly, a very recent single-cell study characterized a common progenitor of cardiomyocytes and epicardium, which differ anatomically and transcriptionally from the FHF and SHF progenitors (Tyser et al. 2021).

4.2 SINGLE-CELL ANALYSIS OF PRE- AND POST-NATAL HUMAN HEARTS

Recent single-cell studies have covered the early heart development of embryonic 4.5 to 9 WPC stages (Sahara et al. 2019; Cui et al. 2019; Asp et al. 2019) and fetal cardiogenesis from 9 to 25 WPC (Sahara et al. 2019; Suryawanshi et al. 2020; Cui et al. 2019). In paper III, we present single-cell analysis of seven pre-natal human hearts ranging from 4.5 to 10 WPC.

Cui and colleagues (Cui et al. 2019) studied human developing hearts ranging from 5 to 25 WPC. They identified different transcriptional profiles between spatial subsets of cardiomyocytes including compact and trabecular as well as atrial and ventricular cardiomyocytes. Furthermore, pseudotemporal analysis of the cardiomyocyte trajectory showed dynamic change in specific gene modules, especially in extracellular matrix-related gene expression. In addition to cardiomyocytes, development of cardiac fibroblasts, valvular and endothelial cells was characterized, and finally the human cell populations were compared with their counterparts in mouse at several developmental time points. Across the

compared cell types, cardiomyocyte transcriptomes showed highest level of conservation between species.

Asp and colleagues (Asp et al. 2019) studied 4.5-5, 6.5 and 9 WPC human embryos using spatial transcriptomics approach and 6.5-7 WPC embryos using single-cell RNA sequencing. Compared to the study by Cui and colleagues, this study has fewer analyzed time points, but the annotation of cell types, especially different cardiac fibroblast populations, is more detailed. Combination of single-cell and spatial transcriptomics allowed Asp and colleagues to specify the spatial subpopulations, for example distinguishing fibroblasts related to small and large vasculature and cardiac skeleton. In addition, a small population of cardiac neural crest cells was identified, which was not reported in other single-cell studies on human embryos. The whole dataset was converted into an online-based resource, which I have utilized in my projects.

Recent single-cell analyses on later pre-natal time points by Suryawanshi and colleagues revealed the cellular composition in the more mature fetal heart. Single-cell analysis of the human fetal hearts and comparison between healthy fetal hearts at 19 to 22 WPC and a heart with congenital heart block at 21 WPC indicated some differences in transcriptional signatures, most notably in interferon-stimulated genes and genes related to fibrosis (Suryawanshi et al. 2020). As shown by Cui and colleagues, expression of extracellular matrix genes is an integral part of normal heart development, whereas it can also be a hallmark of CHD, as shown by Suryawanshi and colleagues.

Compared to pre-natal human heart, the cellular composition of the adult human heart has been less studied using single-cell RNA sequencing. One reason for the lack of single-cell studies on human adult heart is that the currently popular high throughput microfluidics-based single-cell platforms do not support cells as large as adult human cardiomyocytes. Tucker and colleagues (Tucker et al. 2020) and Litviňuková and colleagues (Litviňuková et al. 2020) circumvented this issue by isolating single cardiac nuclei, and both groups sequenced around 300 000 nuclear transcriptomes each. These studies together provide a cell type atlas of the normal adult human heart, which is useful for understanding mechanisms behind adult-onset heart disease. Furthermore, Wang and colleagues (Wang et al. 2020) used single-cell sequencing to compare cardiomyocytes and non-cardiomyocytes in healthy, failed and recovering adult human hearts, and looked at differential cell composition, homeostasis and genetic networks in the different stages of cardiac disease.

4.3 SINGLE-CELL ANALYSIS OF *IN VITRO* CARDIAC DIFFERENTIATION

As shown in earlier clonal analyses and recent single-cell studies (Bu et al, 2009; Ruan et al. 2019; Friedman et al. 2018; Churko et al. 2018), *in vitro* cardiac differentiation generates both cardiomyocyte and non-cardiomyocyte lineages (Figure 3). Hence, interactions and developmental relationships between these cell types during early development can be studied *in vitro*, which is greatly facilitated by single-cell RNA sequencing. A few labs, including ours, have employed the small molecule mediated Wnt-modulation based cardiac

differentiation of hPSCs for single-cell analysis of *in vitro* cardiogenesis (Ruan et al. 2019; Friedman et al. 2018; Churko et al. 2018; Lian et al. 2012).

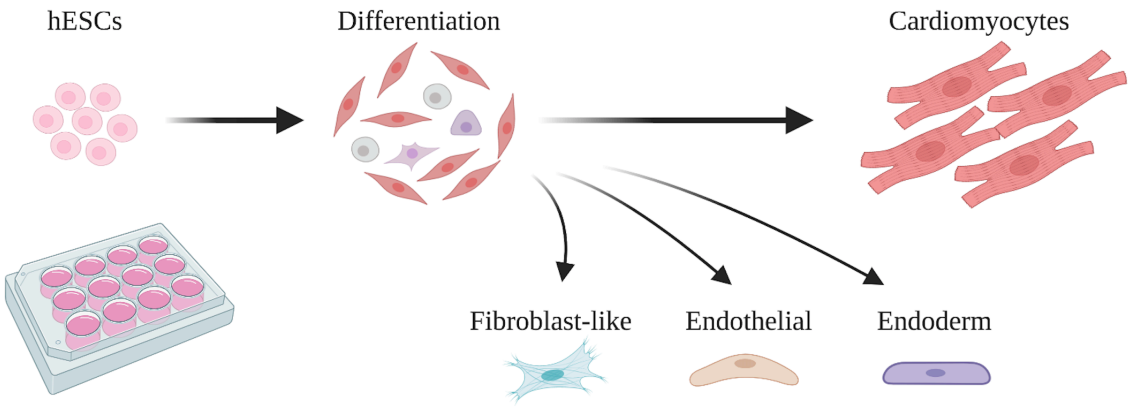

Figure 3. Cardiac *in vitro* differentiation generates cardiomyocytes and other cell types. Created with BioRender.com.

The study by Friedman and colleagues (Friedman et al. 2018) was the first one to analyze multiple time points of cardiac differentiation based on this protocol in single-cell resolution. They identified cardiac progenitors and cardiomyocytes, but also endoderm and outflow tract-like cardiac fibroblasts. The study further identified maturation and hypertrophy-related genetic programs down-regulated in the *in vitro* differentiation system, among which HOPX was highlighted and further validated.

Focusing on transcription factor hierarchies, Churko and colleagues (Churko et al. 2018) sequenced cells from days 0, 5, 14 and 45 of the cardiac differentiation and identified a multitude of cell types, including both atrial and ventricular cardiomyocytes, endoderm, ectoderm and neural crest-like cells. The study suggests that the three transcription factors, TBX5, NR2F2 and HEY2, regulate the electrophysiological properties of different cardiomyocyte subtypes found in the *in vitro* differentiation.

Ruan and colleagues (Ruan et al. 2019) sequenced cells from time points 2, 5, 9 and 14, and focused on the signaling between cardiomyocytes and the endoderm cells found in the *in vitro* cardiac differentiation. Based on differential gene expression in cardiac progenitors, Ruan and colleagues concluded that ETS1 is an important transcription factor regulating cardiac differentiation and dependent on the crosstalk between cardiac progenitors and endoderm. However, ETS1 is only expressed by endothelial cells in our single-cell dataset, which highlights the differences between different datasets and data interpretation.

To conclude, the recent single-cell RNA sequencing studies by us and other laboratories have clarified the cellular heterogeneity and developmental paths of the hPSC-based cardiac differentiation. These new insights are important steps towards comprehensive understanding of the early cardiac progenitor differentiation, which cannot be easily accessed and studied in human embryos.

5 RESEARCH AIMS

The overall aim of this thesis is to gain new insights into the development of early cardiac progenitors, and to apply this knowledge to advance the understanding of CHD pathogenesis.

The specific aims of the constituent papers are as follows:

Paper I: To build a roadmap of cardiac progenitors through single-cell analysis of hESC to cardiomyocyte differentiation, identify human fetal counterparts of myogenic and non-myogenic populations arising during the differentiation, and to investigate the role of ISL1 in the development of these cell populations.

Paper II: To map the genetic landscape of non-syndromic Tetralogy of Fallot by whole-genome sequencing, and to use integrative analysis to pinpoint enrichment of disease-associated gene expression among the single-cell transcriptomes of early cardiac cells.

Paper III: To comprehensively interrogate cardiogenesis of hESCs and human embryonic/fetal heart by single-cell analysis, and thereby understand the cellular origins of certain congenital cardiac malformations in human.

6 METHODOLOGY

6.1 HUMAN PLURIPOTENT STEM CELLS

The hPSCs, which are often used in *in vitro* differentiation protocols, include 1) hESCs, derivatives of the inner cell mass of human blastocysts (Thomson et al. 1998, Evans and Kaufman 1981, Martin 1981), and 2) human induced pluripotent stem cells (hiPSCs), derived from somatic cells by direct reprogramming through expression of transcription factors *OCT3/4*, *SOX2*, *MYC* and *KLF4* (Takahashi et al. 2007; Yu et al. 2007). Both types of hPSCs represent a pre-gastrulation stage of development, are capable of unlimited proliferation in an undifferentiated state, and are capable of generating derivatives of the three germ layers (Zhu and Huangfu 2013, Yu et al. 2007, Takahashi et al. 2007, Thomson et al. 1998). Maintenance of both types of hPSCs in an undifferentiated state requires either feeder cells, usually mouse embryonic fibroblasts, or a coating mimicking the extracellular matrix, such as Matrigel. Both of these alternatives contain xenogeneic components, and as a better alternative completely synthetic coating materials have been developed, allowing the use of hPSC derivatives in regenerative medicine applications (Villa-Diaz et al. 2013).

One inherent challenge of hESCs and hiPSCs is the heterogeneity between cell lines, and their potentials to differentiate to different cell types (Hu et al. 2010; Osafune et al. 2008). For example, the hESC line ES08 is ideal for pancreatic differentiation, whereas ES03 is preferred for cardiac differentiation (Osafune et al. 2008). Heterogeneity between hiPSC lines is especially pronounced due to epigenetic memory of the tissue of origin, as well as the epigenetic and genetic scars generated during the direct reprogramming (Kim et al. 2010; Polo et al. 2010; Gore et al. 2011; Laurent et al. 2011; Lister et al. 2010). Consequently, hESCs may be a more robust source for directed differentiation. This opinion might be biased, and hiPSC have certain advantages, such as usage of patient-derived cells to study disease phenotypes.

6.2 FROM STEM CELLS TO BEATING CARDIOMYOCYTES

One of the main premises of my thesis is that hPSC cardiac differentiation recapitulates many aspects of early heart development ‘in a dish’. The *in vitro* cardiac differentiation protocols are designed to follow the signaling regulating lineage differentiation in an embryo, and the genetic programs taking place in hPSC cardiac differentiation follow a similar sequence *in vitro* as they do *in vivo* (Später et al. 2014; Mummery et al. 2012).

The first achieved *in vitro* cardiac differentiations of hPSCs were based on embryoid bodies, spontaneously forming aggregates of hPSCs which can be directed towards different lineages by adding specific cues to the culture medium (Schuldiner et al. 2000; Kehat et al. 2001; Itskovitz-Eldor et al. 2000). However, embryoid bodies generate less than 10% cardiomyocytes (Kehat et al. 2001), and to improve the yield of cardiomyocytes, various protocols were developed adding for example mouse visceral endoderm-like cell conditioned medium (Xu et al. 2008) and exogenous defined growth factors mimicking mesoderm

inducing signals (Burridge et al. 2012; Kattman et al. 2011; Yang et al. 2008). Another approach for cardiac differentiation relied on co-culture with endodermal cells (Mummery et al. 2012).

Monolayer-based approaches were developed as an alternative to the embryoid body-based cardiac differentiation, providing more defined differentiation conditions and higher cardiomyocyte yield. These protocols are typically based on addition of specific developmental cues at specific timepoints to the cell culture system to first direct the hPSCs towards one of the germ layers and thereafter specify the lineage. In growth factor-based method, activin A, bone morphogenic factor (BMP4), FGF2, VEGF and Dickkopf-1 (DKK-1) are first added in RPMI/B27 medium to induce the formation of mesendoderm (Schuldiner et al. 2000; D'Amour et al. 2005). Thereafter, inhibition of canonical Wnt pathway directs the differentiation towards cardiac mesoderm (Lian et al. 2012).

Small molecule-based modulation of canonical Wnt signaling pathway was introduced using GSK3 inhibitors with small-molecule inhibitors of Wnt ligand production (IWPs)/PORCN inhibitors (Burridge et al. 2012; Lian et al. 2013; Lian et al. 2012; Foo et al. 2018). This protocol is based on the notion that endogenous Wnt/β-catenin pathway is required for cardiac differentiation from hESCs (Paige et al. 2010; Lian et al. 2013). During embryonic development, Wnt signaling initially promotes cardiac differentiation, but later inhibits the progenitors from being differentiated into the cardiac derivatives (Naito et al. 2006; Kwon et al. 2007). Accordingly, an initial addition of GSK3 inhibitor allows the accumulation of β-catenin into the cytoplasm and nucleus, where it acts as a coactivator of transcription factors in the TCF/LEF family, and activation of downstream targets, such as *BRACHYURY*/*T*, leading to induction of mesoderm formation (Showell et al. 2004). In the absence of the inhibitor, β-catenin is targeted by a destruction complex, which GSK3 is a part of, to ubiquitination and degradation by the proteosome. Alike the growth factor method, subsequent inhibition of Wnt ligand production by IWPs directs the differentiation towards cardiac mesoderm (Lian et al. 2013).

Following these protocols, the differentiating cells follow genetic programs familiar from the *in vivo* cardiac differentiation. The cells stay in a pluripotent-like state expressing pluripotency markers *OCT4* and *SOX2* until they start expressing *T*, *MIXL1* and *MESP1* related to mesoderm specification, followed by the sequence of *ISL1*, *GATA4*, *NKX2-5* and *TNNT2* expression, corresponding to cardiac progenitors, specified cardiac progenitors and early cardiomyocytes (Protze et al. 2019; Mummery et al. 2012; Kehat et al. 2001).

When differentiated, both *in vivo* and hPSC-derived cardiomyocytes 1) contract through actions by sarcomere and calcium supply from the sarcoplasmic reticulum, 2) spontaneously depolarize through membrane-bound ion pumps and channels, and 3) are electrically coupled with other cardiomyocytes via gap junctions (Jiang et al. 2018; Foo et al. 2018; Parikh et al. 2017). However, the cardiac differentiation of hPSCs generates cardiomyocytes that are functionally and structurally different from adult cardiomyocytes (Jiang et al. 2018). The differences between *in vitro*-derived and human adult cardiomyocytes include size,

complexity of the t-tubule network, number of nuclei, gene expression, bioenergetics and electrophysiological properties, the *in vitro*-derived cardiomyocytes being closer to the embryonic- and fetal-like phenotype in these aspects (Robertson et al. 2013). In our study presented in paper I, we found that the day 15 hESC-derived immature cardiomyocytes resemble the 5-week human embryonic ventricular cardiomyocytes at single-cell level (Paper I, Figure 2).

The choice of differentiation method can be based on several criteria, such as atrial versus ventricular phenotype. The small molecule modulation of canonical Wnt pathway has been shown to more robustly generate cardiomyocytes with the ventricular phenotype compared to the growth factor-based protocols (Lian et al. 2012). Furthermore, other protocols have been developed to differentiate hPSCs into atrial cardiomyocytes and pacemaker cells (Laksman et al. 2017; Lee et al. 2017; Protze et al. 2017), epicardial cells (Guadix et al. 2017), cardiac fibroblasts (Zhang et al. 2019), as well as endothelial and vascular smooth muscle cells (Patsch et al. 2015; Palpant et al. 2017).

6.3 GENE EDITING

Gene editing with minimal off-target effects is essential for the utilization of stem cells in studying development and congenital defects. In my current project, which has not yet reached the maturity to be included in this thesis, I study the role of one of the candidate genes identified in a Tetralogy of Fallot patient cohort in heart development through genetic manipulation of hESCs. In addition, we used genetic ablation of *ISL1* expression to study its role in the differentiation of myogenic and non-myogenic progenitors in the *in vitro* cardiac differentiation system.

Homologous recombination (HR), zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) have been widely used for gene editing in both *in vitro* and *in vivo* systems in academia and industry (Gordon et al. 1980; Smithies et al. 1985; Thomas et al. 1987; Kim et al. 1996; Urnov et al. 2010; Hockemeyer et al. 2011, Cong et al. 2013). Although important discoveries have been made using these methods, they involve a lengthy and complicated design, difficulty to avoid off-target effects and large expression vectors (Gupta and Musunuru 2014; Miller et al. 2011).

Nowadays, the most popular method for precise gene editing is based on clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated nuclease (Cas9) (Mali et al. 2013). Due to the great impact on genetic research and other fields, Emmanuelle Charpentier and Jennifer Doudna were awarded the 2020 Nobel prize in chemistry for their discovery of the CRISPR/Cas9 system. CRISPR/Cas9 is based on the adaptive immune system of bacteria and archaea, which utilizes Cas proteins in combination with guide RNAs. The guide RNAs are expressed from protospacers, foreign DNA sequences, which bacteria and archaea collect and incorporate in their own genomes. These foreign sequences are expressed as short RNAs with adapter sequences, which are recognized by the Cas protein

and bind to foreign DNA sequences, such as those of bacteriophages. Upon guided binding, the Cas protein catalyzes a cleavage and introduces a double strand break. (Jinek et al. 2012)

ZFNs, TALENs and CRISPR/Cas9 systems introduce double strand breaks, which the DNA repair system mends by either nonhomologous end joining (NHEJ) or in the presence of a DNA template by homology-directed repair (HDR). NHEJ is an error prone process leaving small mutations often leading to a frame shift, which are useful for generation of gene knockouts. (Gupta and Musunuru 2014) On the other hand, HDR can be utilized to precisely induce specific mutations into the DNA sequence, which is often used to correct mutations in patient-derived hiPSC lines and knocking in reporter alleles (Ben Jehuda et al. 2018). The most commonly used Cas protein, Cas9, recognizes a sequence upstream to -NGG sequence, a protospacer adjacent motif (PAM). Mali et al. demonstrated bioinformatically, that there are approximately 190,000 possible target sequences, covering 40.5% of the human genome (Mali et al. 2013). Therefore, researchers are given a great freedom for designing gene edits.

6.4 SINGLE-CELL RNA SEQUENCING

If CRISPR/Cas9 is the state-of-the-art method for gene editing, single-cell RNA sequencing is the state-of-the-art method for gene expression analysis in cells and tissues. Brady et al. (1990) and Eberwine et al. (1992) published the pioneering work towards the development of the single-cell RNA sequencing technology by generating complementary DNA (cDNA) libraries from cells isolated from mouse bone marrow and rat hippocampus, respectively. The early single-cell RNA sequencing was based on high-density DNA microarray chips (Hwang et al. 2018), and the next-generation sequencing-based single-cell RNA sequencing first appeared in 2009, when it was applied to analyze the transcriptome of a single mouse blastomere (Tang et al. 2009). Li et al. (2016) and Kokkinopoulos et al. (2015) were among the first to report a single-cell analysis of murine heart development, revealing differences between *in vitro* and *in vivo* cardiac development.

6.4.1 Single-cell library preparation and sequencing

The single-cell RNA sequencing workflow (Figure 4) starts with isolation of single cells into a hypotonic lysis buffer, which can be carried out using various methods. Common single-cell isolation methods include 1) dilution-based pipetting of cells in low concentration suspensions, 2) microscope-guided capillary pipet cell picking, 3) fluorescence-associated cell sorting (FACS), 4) laser capture microdissection for isolation of cells from solid tissues, 5) microfluidics, such as that used in Drop-seq, Fluidigm and 10x Genomics workflows and 5) nanowell-based systems (Hwang et al. 2018, Shapiro et al. 2013).

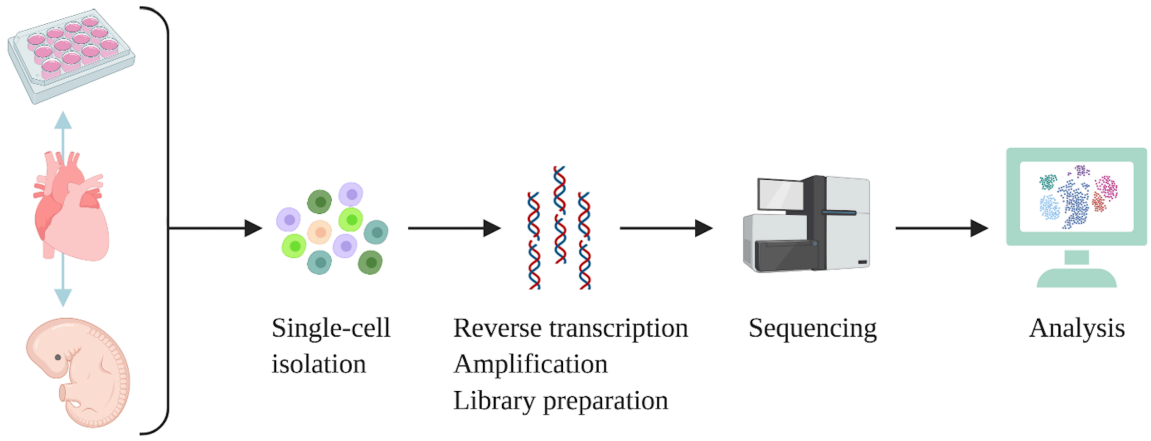

Figure 4. General workflow for studying heart development at single-cell resolution. Created with BioRender.com.

After lysis, the single cells are processed into cDNA libraries by reverse transcription of first-strand cDNA using poly (dT) primers, which anneal to the poly(A) tails of eukaryotic mRNAs; second-strand synthesis using an engineered reverse transcriptase; and cDNA amplification by either PCR or *in vitro* transcription (Hwang et al. 2018). These steps are common for the commonly used protocols for single-cell library preparation (Ziegenhain et al. 2017). Choice of methods for cell isolation and cDNA library preparation depends on the desired cell number, cost, read depth, throughput, and acceptable level of bias (Hwang et al. 2018, Shapiro et al. 2013). Currently, the Smart-seq3 workflow reconstructs the single-cell transcriptome in the highest detail by generating full-length high-depth libraries, i.e. generates more reads and detected genes per cell, and is therefore applicable for splice variant analysis (Picelli et al. 2014, Ziegenhain et al. 2017). On the other hand, the microfluidics-based platforms offer a higher throughput and cost-efficiency (Bageritz and Raddi 2019, Ziegenhain et al. 2017).

The Illumina platform is based on sequencing by synthesis chemistry, and is widely used for the sequencing step. The choice of Illumina next-generation sequencing system is largely dependent on the scale of the single-cell study, which determines the reasonable depth of sequencing. The acceptable sequencing depth, depending on the experimental plan, can vary from 10,000 to several million reads per cell. In addition to the library preparation method, the sequencing depth can be further affected by choosing either paired-end or single-read sequencing, which determines whether the DNA is sequenced from one or both ends. Furthermore, during the library preparation, either single- or dual-indexes can be added to the cDNA library, which allows multiplexing, i.e. multiple libraries being pooled and sequenced together. (Single-Cell Sequencing Workflow: Critical Steps and Considerations, Illumina 2020)

6.4.2 Principles of basic single-cell analysis

The technological breakthroughs leading to the generation of detailed large-scale single-cell data bring along new challenges for developing analytical tools to fully utilize it. The data

analysis can be carried out using in-house pipelines or by following available workflows, such as Seurat (Butler et al. 2018).

After sequencing, data analysis starts with file conversion and demultiplexing (if multiplexing was applied), mapping reads to a reference genome, data quality control and filtering, normalization, and finally downstream analyses, such as data visualizations by clustering analysis (Gao 2018). The quality control and filtering out poor quality cells and doublets are commonly based on library size, number of genes per cell and proportion of mitochondrial genes; too large library indicating doublets and high mitochondrial gene expression commonly thought to denote apoptotic or pro-apoptotic cells (Ziegenhain et al. 2017). Furthermore, the mRNA content per cell is usually normalized using either External RNA Control Consortium (ERCC) spike-ins or unique molecular identifiers (UMIs) method, mitigating the technical noise (Gao 2018). Other sources of technical variation include batch effect and dropout amplification bias, whereas cell cycle phase is an example of confounding biological variation (Hwang et al. 2018). What comes to cell cycle phase, it can contribute to meaningful and interesting biological variation, but can also mask other features of the dataset, such as continuous differentiation axes, as demonstrate in paper I.

When the confounding noise has been addressed, exploration of the biological variation can start. Usually, one would start with principal component analysis for dimension reduction (Wold et al. 1987), which can also be used to explore sources of biological variance in the dataset. Identification of cell populations is one of the central goals of single-cell analysis, and is achieved by utilizing machine learning-based unsupervised clustering algorithms. Popular unsupervised clustering algorithms include k-means, hierarchical, density-based and graph-based clustering using the Louvain algorithm (Andrews and Hemberg 2018). These algorithms group single-cells in a multidimensional space based on their transcriptional signatures. Groups of cells and relationships between them are visualized by a dimensional reduction method, commonly using either t-distributed stochastic neighbor embedding (t-SNE) or uniform manifold approximation and projection (UMAP). Compared to t-SNE and other dimensionality reduction techniques for single-cell analysis, UMAP has been shown to better conserve the global data structure, i.e. the relationship between clusters, and is particularly useful for the visualization of continuous cell transitions and branch points (Becht et al. 2018). The clustering can be further utilized for marker gene calling, which can be inputted into gene ontology (Ashburner et al. 2000) and pathway analysis for identification of cell types and biological processes.

6.4.3 Developmental trajectories and co-regulated gene modules

Reporter alleles have been long used to trace specific lineages and subpopulations based on marker gene expression in the heart and other organ systems. Single-cell RNA sequencing has enabled the development of computational analysis methods for identification of small subpopulations, lineages and population-specific gene regulatory networks. In the work included in my thesis, we explored and utilized some of the available algorithms to elucidate specific progenitor behaviors and dynamics along cardiac differentiation.

Although single-cell RNA sequencing only captures temporal snapshots of lineage development, several methods have been developed to trace the origin and future destiny of individual cells. RNA velocity is a method for the interference of differentiation direction of individual cells (Svensson and Pachter 2018; La Manno et al. 2018). RNA velocity is based on the distinction between spliced and unspliced RNA species, which can be derived from the single-cell RNA sequencing data. We applied RNA velocity analysis in our study presented in paper I, which revealed new insights regarding the developmental dynamics along the cardiac progenitor differentiation.

Trajectory analysis is a one of the most popular approaches for single-cell analysis of lineage differentiation. Building a differentiation trajectory for lineage interference requires single-cell analysis of multiple adjacent time points, the assumption being that the cells at slightly different developmental stages build a continuous spectrum. Multiple approaches have been developed for building developmental trajectories, including Monocle (Cao et al. 2019), Waterfall (Shin et al. 2015), SCUBA (Marco et al. 2014), DPT (Haghverdi et al. 2016), TSCAN (Ji and Ji 2016) and Slingshot (Street et al. 2018). The main goal of these methods is to reconstruct pseudotime by computationally ordering the individual cells along trajectories that can either be linear or branched. Pseudotime should on some level correlate with the data time points, but given that the cells captured at one time point are heterogeneous in their level of differentiation, pseudotime can better order the cells based on their transcriptional profiles.

In addition to differentiating lineages, single-cell analysis provides information on genes that are co-expressed or co-regulated across groups of cells and pseudotime. Reconstruction of gene regulatory networks has been used for bulk RNA sequencing data, while single-cell RNA sequencing data can be analyzed using similar methods at higher cellular resolution. The basic assumption with gene module analysis is that groups of genes that are highly co-expressed across single cells may also be co-regulated. (Chen et al. 2019) In the study presented in paper II, we utilized Louvain community analysis algorithm (Traag et al. 2019) in the Monocle3 pipeline to find modules of co-expressed genes among the genes that change as a function of Monocle-derived pseudotime (Cao et al. 2019; Qiu et al. 2017; Trapnell et al. 2014). We identified modules of co-regulated genes that were specific for specific stages of cardiac differentiation while some modules were present across multiple stages. Other pipelines providing a framework for similar analysis include WGCNA (Langfelder and Horvath 2008; Li et al. 2018) and SCENIC (Aibar et al. 2017).

Various pipelines have been developed for the different classes of single-cell analysis, many more than presented here. The pipelines are becoming increasingly user friendly, and several meta-analyses have compared their robustness, allowing more wet-lab researchers such as myself to utilize them efficiently. Many of the pipelines are complementary to each other, while others are better than others, and despite the positive impact of new pipelines, the choice of best single-cell workflow is not an obvious one.

6.5 IDENTIFICATION OF DAMAGING GENETIC VARIANTS

A number of genetic approaches has been developed to identify genetic variation on individual and population level. These approaches include single nucleotide polymorphism (SNP) arrays, copy number variant platforms and Sanger sequencing, as well as next-generation sequencing platforms whole-exome and whole-genome sequencing (Fahed et al. 2013). These genetic tools have been combined with different study designs including: 1) family studies, in which inherited mutations are tracked by sequencing and phenotyping multiple family members, and 2) cohort studies, in which a large number of unrelated cases and healthy controls undergo phenotyping and sequencing for identification of rare genetic variants (Williams et al. 2019).

Studies aiming to identify novel genetic variants associated with a patient phenotype utilize next-generation sequencing rather than the conventional sequencing platforms. In particular, a third type of study design, trio studies, in which the entire exomes or genomes of a proband and unaffected parents are sequenced, has enabled the detection of novel damaging *de novo* and recessive variants and variants with unclear Mendelian inheritance (Shabana et al. 2020). Due to the development of increasingly cost-effective methods for library preparation and sequencing, an increasing number of patients with various cardiac defects have been included in comprehensive genetic evaluation (Nees and Chung 2020). As a result, large sequencing studies have shed light on the prevalence and architecture of rare disease-causing mutations in CHD patients (Page et al. 2019, Jin et al. 2017, Sifrim et al. 2016, Preuss et al. 2016).

In the sequencing study presented in paper II, we performed whole-genome sequencing of pediatric cardiac patients and their parents, and here I will give a brief overview on the technical workflow for variant calling. Similar to single-cell RNA sequencing, the next-generation sequencing raw reads are first mapped to the reference human genome. Genome Analysis Toolkit (GATK) provides a comprehensive framework for variant analysis and was used by our collaborators to identify damaging genetic mutations in the patients (Van der Auwera et al. 2013; McKenna et al. 2010). In brief, analysis of the variants involved quality controls, as well as calls for single nucleotide variants, indels and copy number variants. After the genetic variants are identified, they are annotated according to their pathogenicity and frequency in the general population. How damaging the identified *de novo* and recessive mutations are, and whether they cause loss of function, is determined using *in silico* prediction algorithms, including MetaSVM and CADD. The variants are further prioritized according to predictions on their effect on protein function, and whether they are synonymous, non-synonymous, missense or nonsense, cause a frameshift, or cause a change within essential splice, start or stop site. After variant calling and prioritization, one can perform a burden analysis of damaging variants in the patient population as well as a functional enrichment analysis using similar methods as one would use in a single-cell analysis, i.e. gene ontology and pathway analysis. (Van der Auwera et al. 2013; McKenna et al. 2010)

6.6 STATISTICAL ENRICHMENT AND CROSS-REFERENCING

After mapping the genetic landscape of a disease or phenotype, such as tetralogy of Fallot, in a large patient cohort, it can be interesting to understand which cell types these damaging variants affect. Both developmental and acquired diseases change the composition of cells and cell homeostasis in the affected tissue, and single-cell RNA sequencing data can provide information on even small changes in small and large cell populations and the gene regulatory networks within cells (Paik et al. 2020). Furthermore, information on the cellular composition of healthy tissues and their single-cell transcriptional profiles can aid in pinpointing cell types most affected in specific pathological conditions.

To intersect single-cell transcriptomics with gene hits from whole genome sequencing, we used statistical tests based on Poisson and hypergeometric distribution in the study presented in paper II. In these analyses, we cross-referenced the *de novo* and recessive damaging genetic variants identified in Tetralogy of Fallot patients with 1) gene ontology terms, 2) marker genes for specific *in vivo* and *in vitro* cardiac populations identified through differential gene expression analysis, and 3) gene regulatory networks identified by module-based co-regulation analysis in the human embryonic stem cell *in vitro* cardiac differentiation.

Hypergeometric test is similar to the binomial test, and the popular enrichment analysis tools for gene ontology and pathway analysis, such as Metascape, Panther and String, are based on these statistical tests. Both binomial and hypergeometric tests can be used to calculate whether genes with damaging variants are over- or underrepresented in lists of genes derived from single-cell data, such as population markers. For example, Li and colleagues used hypergeometric test to study the enrichment of neuropsychiatric risk genes in modules of co-expressed genes derived from single-cell sequencing of pre- and post-natal human brains (Li et al. 2018). To follow the benchmark set by Li and colleagues, we used the hypergeometric distribution method to calculate the enrichment of Tetralogy of Fallot candidate genes within modules of co-regulated genes. Lastly, to calculate the statistical enrichment of damaging *de novo* variants within population marker genes, we used Poisson distribution-based denovolyzeR statistical tool, which takes into account the expected number of *de novo* variants in a given population, mutability of the gene and the number of study subjects (Ware et al. 2015).

6.7 ETHICAL CONSIDERATIONS

Human embryonic stem cells: The generation of hESC lines involves destruction of human blastocysts, and thereby stem cell research can be seen as ethically controversial. The blastocyst, an early stage embryo, used for this purpose is usually obtained from an *in vitro* fertilization clinic, in which couples can give an informed consent to whether they allow the use of excess blastocysts not to be implanted to be used for scientific purposes (King and Perrin 2014). An alternative for embryonic stem cells is hiPSCs, derived by reprogramming somatic cells to become pluripotent. As the generation of iPSCs does not depend on human

embryos as a source, they are not seen controversial, as long as the donors have given their consent. The prominent ethical problem with hESC cell research is the moral status of the blastocyst that was destroyed in order to obtain the embryonic stem cells line. If the embryo should be treated as a human and given a moral status, generation of hESC lines is not justified, as we cannot obtain an informed consent from the embryo. One could argue, on the other hand, that these blastocysts, although potential human beings, would not be used for *in vitro* fertilization, and would anyway be disposed. The third viewpoint is a utilitarian one, namely that despite our view to the moral status of the embryo, the knowledge obtained from stem cell research justifies destroying human embryos.

Collection and handling of patient information: Part of my PhD project is focused on studying mutations identified in a cohort of CHD patients. The first ethical issue regarding this study is that the patients are infants, and thereby not capable of giving an informed consent by themselves. Instead, the informed consent is obtained from the parents, who receive information concerning the study and its risks. The study itself does not cause harm for the patients, and only requires a blood sample for extraction of genetic material. However, the genetic information that can be extracted from the simple blood samples can reveal sensitive information about the study participants, such as predisposition to certain health problems. The second ethical issue concerns the handling of patient information. The most important concern for us is to keep the patient information confidential so that it will not become available for anyone else except the researchers working on the project. One way to ensure this was to encode the data for transport between different institutions involved in the study so that it is not easily accessible for someone outside academic research. Another way to protect sensitive patient information is to assign each patient an ID number, which will be used during the analysis to link together the patient phenotype and genotype instead of using the patients' names or social security numbers. Lastly, the patients will be anonymous in the planned publication.

Prenatal human cardiac samples: Our studies have involved obtaining human embryonic tissues for single-cell analysis as well as immunofluorescence and confocal microscopy-based analyses of markers for cardiac populations. The human embryonic tissues were obtained from a clinic at Karolinska Hospital following a valid ethical permit. When possible, we also used single-cell analyses done by other laboratories in order to avoid redundancy and use the available single-cell data to its full potential.

7 RESULTS

7.1 PAPER I

With this study, we aimed to deeply characterize the developmental path of cardiac progenitors through single-cell transcriptional analysis of hESC cardiac differentiation. Generation of high-quality Smart-seq2 single-cell cDNA libraries from cells collected at closely adjacent differentiation time points allowed us to build a connected roadmap of the cardiac differentiation process.

A widely adopted WNT-modulation based protocol (Lian et al. 2013) was used to achieve cardiac differentiation of hESCs. In order to capture a broad spectrum of cellular states, cells were collected on days 3-9 and 15, each time point in two separate differentiation batches. These timepoints cover the differentiation process from undifferentiated cells expressing pluripotency-associated genes to beating cardiomyocytes expressing canonical cardiomyocyte marker genes. In order to avoid collection bias, cells were picked manually using a hand-pulled capillary needle. In total, 1024 single-cell cDNA libraries were prepared for sequencing following the Smart-seq2 protocol. After sequencing and quality control, 925 cells passed the quality metrics and were included in downstream analyses. Despite the relatively small number of sequenced cells, visualization of connected lineage trajectories was possible due to the good sequencing depth and the collection of cells at close timepoints. Seurat version 3 pipeline was utilized for basic analyses, including principal component analysis, dimensional reduction, unsupervised clustering and marker gene analysis. In addition to a differentiation axis of pluripotent stem cells to cardiac progenitors and cardiomyocytes, we observed smaller diverging populations of cells, which were identified as endoderm and cardiac fibroblast-like cells.

In the past years the hESC-derived cardiac cells have been mapped to their human *in vivo* counterparts mostly based on electrophysiological properties and expression of specific genes. A previous study indicated that the hPSC-derived cardiac cells are comparable to human fetal cells at global transcriptional level (Van den Berg et al. 2015). In order to study the single-cell transcriptional resemblance between the hESC-derived populations and cells in a human embryonic heart in more depth, we integrated the *in vitro* data with a previously reported single-cell analysis of a human embryonic heart at 5 WPC (Cui et al. 2019). The cells from the *in vitro* differentiation time points day 3-7 were developmentally too early to transcriptionally overlap with the human embryonic *in vivo* cells, and were therefore excluded from the analysis. The cells from the later time points showed co-clustering with the human fetal heart. Marker gene and gene ontology analyses were used to identify the shared populations, including ventricular cardiomyocytes, atrial cardiomyocytes, fibroblast-like cells likely of epicardial origin, proliferative cardiac myogenic cells, endothelial cells and cardiac progenitors. These results verify the identities of hESC-derived populations, and validate the relevance of the hESC-based differentiation system to study early cardiovascular cells which are difficult to access *in vivo*.

We found that the day-8 and -9 hESC-derived cardiac progenitors clustered together with 5-week embryonic atrial cardiomyocytes, whereas the day-15 hESC-derived cardiomyocytes resembled the embryonic ventricular cardiomyocytes. A previous study comparing hPSC-derived with human fetal cardiac cells found that the *in vitro* differentiation protocol generates cells with both atrial and ventricular phenotypes (Van den Berg et al. 2015). Our results indicate, that these phenotypic differences are at least partially related to the maturity of the cardiomyocytes in the hESC differentiation system. We identified transforming growth factor β-induced protein (TGFBI) as a marker for the cardiac fibroblast-like cells both in the human embryonic heart and the hESC-cardiac differentiation, and validated its expression by immunocytochemistry both *in vitro* and *in vivo*. TGFBI is important in the modulation of the extracellular matrix (Hashimoto et al. 2013), and is likely involved in the migration of the fibroblasts in the developing heart. A cardiac fibroblast-like population has been previously identified in the hPSC-cardiac differentiation (Zhang et al. 2019; Friedman et al. 2018). Our results further indicate, that this population closely resembles epicardium-derived fibroblasts.

Using the UMAP "differentiation roadmap", we identified the branching points of cardiac differentiation where the non-myogenic cell types diverge from the cardiomyocyte lineage. Based on this analysis, the cardiac fibroblast-like cells do not originate from the same early cardiac progenitor population as the cardiomyocytes. In line with *in vivo* embryonic development (Wei and Wang 2018), the endoderm population diverges from the cardiac differentiation at a branch point population, which expressed mesendoderm markers, such as *T*, *EOMES* and *MIXL1*. Regarding the cardiomyocyte differentiation axis, we identified a branch point between the early progenitors and the committed *NKX2-5*-expressing cardiac progenitors, which also gives rise to a small population of endothelial cells. We found, that this branch point was characterized by the expression of specific extracellular matrix-related genes, *COL3A1* in particular, suggesting that extracellular matrix plays an important role in cardiac development. This finding is in line with the human developmental *in vivo* data, which has shown that the expression of extracellular matrix proteins, such as *COL3A1*, follow dynamic changes along the cardiomyocyte differentiation (Cui et al. 2019). To further define the branching point, we regressed out the majority of the cell-cycle variance from the data, which further refined the connections between populations, and studied the specific signaling pathways regulating the lineage transition using pathway analysis. Specifically, we found that WNT and integrin signaling-related genes were expressed by the branch point population, which further supports the importance of extracellular matrix-related genes in cardiac differentiation.

Although the dimensional reduction is an excellent tool for studying the global data structure and the transcriptional differences, similarities and connections between populations, it does not fully utilize the transcriptional information hidden in the single-cell RNA sequencing data. Therefore, we employed RNA velocity analysis, which uses the information on spliced and unspliced mRNAs, to derive the developmental direction of each cell (La Manno et al. 2018; Svensson and Pachter 2018). Interestingly, RNA velocity vectors indicated that progenitors maintain their undifferentiated state until the abovementioned branching point,

after which the cardiac lineage shows differentiation direction towards cardiomyocytes, indicating commitment to a cardiac fate. Maintenance of stemness is an important property of cardiac progenitors, allowing them to proliferate and maintain a pool of migratory cells in the developing heart before cardiac commitment and differentiation into cardiomyocytes of the cardiac chambers and the outflow tract (Zhou et al. 2011).

To explore the role of ISL1 transcription factor in this differentiation process, we conducted additional single-cell analysis of *ISL1* knockout hESC-derived differentiating cells at day 6 of the cardiac differentiation protocol. We found, that loss of *ISL1* expression did not significantly affect the transcriptional signature of the day-6 progenitors. However, similar to a recent study on *ISL1* knockout hPSC cardiac differentiation (Quaranta et al. 2018), we saw a delay in the cardiac commitment of the progenitors, as the *NKX2-5*-expressing population was missing in the knockout differentiation at this timepoint. Similar to wild-type, the *ISL1* knockout hESC generated a small endoderm population. In addition to that, the *ISL1* knockout cells generated an ectopic neuroectoderm population. ISL1 is known to be involved in the differentiation of non-cardiomyocyte lineages, such as neuronal cells (Zhao et al. 2020; Teratani-Ota et al. 2016) and the endoderm-derived pancreatic β-cells (Gupta et al. 2018; Pezzolla et al. 2015). Further research is needed to study the specific role of ISL1 in the fate decision between cardiac and neural progenitor differentiation, and to identify where the neural-like population diverges from the cardiac and other differentiation axes. To gain further insight, single-cell RNA sequencing of *ISL1* knockout differentiation at several timepoints of the early differentiation should be conducted.

In summary, paper I provides novel insights in the differentiation of early cardiac progenitors at single-cell resolution. In particular, by generating a connected roadmap of hESC-based cardiac differentiation, we were able to study the emergence and segregation of cardiac progenitors and lineages from pluripotent cells, and add new detail on the molecular mechanisms regulating these events.

7.2 PAPER II

Previously published large whole-exome sequencing studies have mapped the genetics underlying the pathogenesis of Tetralogy of Fallot in patient cohorts of European descent (Page et al. 2019; Jin et al. 2017). However, little is known about other cohorts. In this study, whole-genome sequencing was performed on 150 Chinese Tetralogy of Fallot cases and their healthy parents following informed consent and ethical approval. In order to exclude syndromic cases, patients were screened for DiGeorge syndrome-associated deletion in the genomic locus 22q11.2. Among the 150 trios, two were found to be biologically unrelated and two of the probands had trisomy 21, a chromosomal aneuploidy commonly associated with Tetralogy of Fallot. These four trios were excluded from the study. This study is focused on covering the genetic findings in the coding genome, whereas the non-coding genetic variants will be reported in a subsequent study.

Thanks to the sequencing of both proband and two unaffected parents, damaging *de novo* mutations could be identified. In total 140 protein-altering *de novo* mutations, including single nucleotide variants and indels and excluding synonymous mutations, were identified in 92 cases. 28 of these findings were loss-of-function mutations. According to a global burden test, the Tetralogy of Fallot cases harbored more damaging *de novo* mutations, particularly loss-of-function mutations, than the general population in average. In addition, a significant enrichment of damaging *de novo* variants was found among human genes that are intolerant to loss-of-function mutations, as well as known CHD susceptibility genes. These results indicate, that *de novo* mutations significantly contribute to the disease risk in the cohort.

In addition to single nucleotide variants and short indels, whole-genome sequencing reliably detects copy number variants. Copy number variants are usually associated with genetic syndromes, but rare *de novo* copy number variants can also cause non-syndromic forms of CHD and Tetralogy of Fallot (Soemedi et al. 2012; Greenway et al. 2009). In total 12 small and 4 large *de novo* copy number variants were identified in 14 patients. Out of these 16 copy number variants, 8 disrupted exonic regions and the remaining 8 were found outside the coding genome. Whereas the copy number variants of the coding region disrupted some of the known cardiac genes, further analysis is needed to fully understand the role of the noncoding ones in the pathogenesis of Tetralogy of Fallot. In particular, whether the noncoding copy number variants disrupt cardiac-specific regulatory regions.

In addition to damaging *de novo* mutations, rare damaging recessive, both homozygous and compound heterozygous, variants have been previously shown to be enriched in CHD cases (Jin et al. 2017). In the Chinese cohort, 28 cases harbored 32 rare damaging recessive (10 homozygous and 12 compound heterozygous) genotypes, when using relatively strict criteria. If the criteria are relaxed, 236 recessive variants could be found.

In order to gain an overview of the biological function of the genes with damaging mutations in the Tetralogy of Fallot cohort, gene ontology enrichment analysis was performed including all damaging *de novo* and recessive variants. 167 significant gene ontology terms were identified, which grouped into 21 enriched clusters. The terms associated with the top significantly enriched clusters were related to cardiac function and development, as well as sensory organ development. The functional enrichment among the Chinese Tetralogy of Fallot cases was compared with a cohort of Tetralogy of Fallot patients with European ancestry reported by the PCGC (Jin et al. 2017). Interestingly, we found differential functional enrichment in *de novo* variants between these two populations, which is likely related to the differences in the genetic background between the two cohorts.

In order to identify human developmental cardiac cell populations with enrichment of disease-associated gene expression in spatial and temporal manner, the gene set of damaging *de novo* variants was cross-referenced with single-cell RNA sequencing data. For this analysis, we used data from three previous studies reporting single-cell RNA sequencing and spatial transcriptomics of human 6.5-7 week embryonic heart (Asp et al. 2019), single-cell analysis of human 4.5-10 week embryonic heart (Sahara et al. 2019), and single-cell analysis

of *in vitro* hESC cardiac differentiation (paper I). We found enrichment of the *de novo* variant genes among marker genes for the myogenic populations across all single-cell datasets, notably those of the cono-ventricular progenitors. The cono-ventricular progenitors will be described in more detail in paper III. Based on the spatial transcriptomics data, significant enrichment was found among the genes specifically expressed in the outflow tract region. We performed the same enrichment analysis on the *de novo* variants identified in Tetralogy of Fallot cases of the PCGC patient cohort, but no enrichment was found.

To pinpoint timepoints along cardiac differentiation affected by the identified mutations, we intersected both *de novo* and rare recessive findings with pseudotime (single-cell) analysis of the hESC cardiac differentiation. Specifically, we cross-referenced the lists of Tetralogy of Fallot-associated genes with gene co-expression modules identified using the Monocle pseudotime trajectory analysis pipeline. Many of these gene co-expression modules showed timepoint-specific patterns, i.e. some of the modules were expressed early in cardiac differentiation (by the pluripotent/pre-cardiac cells), some at the intermediate progenitor stages, and some towards the end of the differentiation (by committed cardiac progenitors/cardiomyocytes). Hence, by cross-referencing the disease-associated genes with the co-expression modules, we could identify groups of Tetralogy of Fallot-associated genes that are co-expressed as well as find their developmental expression patterns along the cardiac differentiation. We also identified the gene co-expression modules related to VEGF and NOTCH signaling pathways. Through these analyses, we identified a small number of gene co-expression modules which were enriched for the Tetralogy of Fallot-associated genes. Two of these modules were expressed early in the differentiation and two were expressed at the late timepoints. Yet, another module, which showed enrichment of Tetralogy of Fallot-associated genes, was specific for VEGF and NOTCH signaling pathways. Interestingly, we found that several of the known Tetralogy of Fallot-associated genes (*TBX1*, *GATA4*, *TBX5*, *NKX2-5*, *FLT4* and *NOTCH1*) belonged to the same co-expression modules that showed enrichment of disease-associated genes identified in the Chinese patient cohort.

Our study maps the genetic landscape of a non-syndromic Tetralogy of Fallot cohort, identifying causative *de novo* and rare recessive variants in the patient population. By functional annotation and comparison with variants identified in a similar cohort of European descent, specific pathways and cellular mechanisms, as well as differences between ethnic groups were identified. Comparisons with single-cell transcriptomics revealed associations of known and novel Tetralogy of Fallot genes with cono-ventricular progenitors and specific gene expression patterns along cardiac differentiation. This analysis resulted in the identification of a large set of genes with currently no known function in heart development. These genes are both strongly associated with complex CHD and expressed at specific timepoints of cardiac differentiation. Thus, they represent a set of interesting candidates for future studies.

7.3 PAPER III

In this study, single-cell transcriptional analysis was applied to study the cell type heterogeneity in human embryonic and fetal hearts as well as hESC cardiac differentiation, which resulted in the identification of a novel subset of progenitors marked by the WNT signal activator LGR5, a leucine-rich repeat-containing G-protein-coupled receptor. The LGR5+ population emerges in the cono-ventricular/proximal outflow tract region around 4-5 WPC in a human embryo, and a similar progenitor population was also found in the hESC-cardiac differentiation at the early progenitor stage. LGR5 has been previously shown to mark several types of adult stem cells, including those of hair follicles, small intestine and colon, in which it has been shown to regulate proliferation (Barker and Clevers 2010). The study presented here investigates the role of LGR5 in early heart development.

Cell populations arising in the hESC cardiac differentiation were characterized and isolated by FACS and population identities were validated by qPCR analysis. The populations included ISL1+PDGFRA+ multipotent cardiac progenitors on day 3 as well as ISL1+ intermediates and ISL1- differentiated cells expressing lineage markers TNNT2 (cardiomyocytes), HCN4 (pacemaker cells), PECAM1 (endothelial cells), and SMMHC (smooth muscle cells) on days 6 and 10, respectively. The isolated cell types and undifferentiated hESCs were prepared into bulk RNA sequencing libraries and sequenced to study the global transcriptional expression on a population level. Notably, LGR5 was identified as a novel marker for the multipotent cardiac progenitors on day 3, and was selected for further investigation.

FACS analysis verified that the developmental stage of the PDGFRA+ cardiogenic mesoderm marked by LGR5 occurs after MESP1+ mesoderm but precedes the peak of ISL1 expression on day 6. In addition, LGR5 expression was detected in the non-mesodermal lineages on day 3, although these cells did not show ISL1 expression. To test the potency of the LGR5+ mesodermal progenitors, a clonal assay was performed on cells collected on day 3. PDGFRA and LGR5 double positive cells as well as single positive and double negative populations were included in the experiment. Single-cell-derived clones were grown and their differentiation potential was tested under three differentiation conditions, namely those promoting cardiomyocyte, smooth muscle cell and endothelial cell differentiation. Both PDGFRA+LGR5+ and PDGFRA+LGR5- showed the capacity to generate all three cardiac lineages, although PDGFRA+LGR5+ cells favored the cardiomyocyte lineage more frequently than PDGFRA+LGR5- cells. These results suggest that while the LGR5+ cells represent a multipotent progenitor, they predominantly differentiate into cardiomyocytes.

To further investigate the LGR5+ cardiac progenitor population, single-cell transcriptional analysis was performed on cells collected on days 3, 6 and 15 of the cardiac differentiation protocol. These cells generated clusters representing stages of differentiation ranging from undifferentiated cells expressing *OCT4* to cardiomyocytes expressing *TNNT2*. Cells expressing *LGR5* were found in two intermediate clusters, and the expression profiles of *LGR5*+ and *LGR5*- cells in these clusters were explored. Interestingly, cells expressing *LGR5*

showed higher expression of the mesodermal cardiogenic markers *MESP1*, *PDGFRA*, *TBX6*, *BMP2* and *BMP4* as well as SHF-related marker genes *ISL1* and *BMP4* compared to the cells not expressing *LGR5*. The single-cell analysis of hESC cardiac differentiation suggest that *LGR5* expression may be related to the SHF progenitor development.

To relate the *in vitro*-derived cardiac populations to their *in vivo* counterparts, cells isolated from altogether seven human embryonic and fetal hearts ranging from 4.5 to 10 WPC were subjected to single-cell sequencing. This analysis revealed small populations of endothelial cells, smooth muscle cells and fibroblasts, which were removed from further analyses due to low abundance. Subsequent analyses on the pooled data revealed ten clusters named cono-ventricular progenitors, free-wall ventricular progenitors, intermediates of outflow tract, intermediates of ventricle, early atria/pace maker cells, late atria, cardiomyocytes expressing extracellular matrix genes, late cardiac mesenchymal cells, cono-ventricular muscle cell and free-wall ventricular muscle cells.

Analysis of differentially expressed genes between embryonic heart-derived single-cell clusters revealed a substantial difference between the cono-ventricular and free-wall ventricular progenitors. While the former population showed specific expression of Wnt signaling-related genes (e.g. *LGR5*, *RSPO3* and *LEF1*), Wnt signaling was rather inhibited in the free-wall ventricular population. Overall, gene ontology analysis of the genes specifically expressed by the cono-ventricular population showed enrichment of terms related to extracellular matrix, heart development, and Wnt and TGF-β pathways. Genes expressed by the free-wall ventricular progenitors also showed enrichment for extracellular matrix and TGF-β pathway, and additionally showed enrichment for genes related to cell adhesion, positive regulation of cell proliferation and PI3K-Akt pathway.

The localization of the LGR5+ISL1+ cells in the proximal outflow tract was validated by immunostaining in 4.5 WPC human embryonic heart sections. A subpopulation of these cells expressed TNNT2, indicating cardiac specification, as well as Ki67, indicating proliferation of the LGR5+ myogenic population. LGR5 expression was low or undetectable in the later human developmental time points studied. To study whether the LGR5+ population is evolutionarily conserved between human and mouse, immunostaining of Lgr5 in mouse heart sections at E9.5-10.5 and lineage tracing using Lgr5-EGFP-IRES-CreERT2 knockin mice together with Rosa26-floxed stop cassette tdTomato reporter mice were carried out. Neither of the approaches resulted in a signal in the proximal outflow tract, indicating absence of the Lgr5+ cardiac progenitor population in the mouse at all stages of cardiogenesis. These results strongly suggest that the LGR5+ population of the outflow tract is human specific.

To further study the functional role of LGR5 during human cardiogensis, the CRISPR/Cas9 system was used to generate an *LGR5* knockout hESC line. After validation, the *LGR5* knockout cells were subjected to cardiac differentiation. As indicated by qPCR analysis, the *LGR5* knockout did not affect the expression of the mesoderm marker *MESP1* on day 3, whereas the expression of *ISL1*, *LEF1* and *TNNT2* was significantly decreased compared to

the wild type on later time points. Analysis on the day 6 population indicated, that the *LGR5* knockout cells generated less ISL1+TNNT2+ cardiomyocyte intermediates whereas other intermediates (pacemaker, smooth muscle cell and endothelial cell) were not affected. On day 14, the number of TNNT2+ cardiomyocytes was lower in *LGR5* knockout hESC differentiation compared to wild type, while the numbers of smooth muscle and endothelial cells slightly increased. Furthermore, flow cytometry analysis showed, that the *LGR5* knockout hESCs generated approximately one third of the number of cardiomyocytes compared to wild type differentiation. The results were validated by subjecting the *LGR5* knockout and wild type hESCs to an alternative growth factor-based differentiation protocol, which indicated that *LGR5* knockout negatively affects cardiomyocyte differentiation, independent of which protocol is used. Moreover, qPCR analysis on day 14 showed that the *LGR5* knockout cardiomyocytes expressed less cono-ventricular myocyte-specific genes, whereas they expressed specific free-wall ventricular genes on the same level as wild type cardiomyocytes. These results indicate, that LGR5 plays a specific role in the differentiation of the cono-ventricular cardiomyocytes.

Lastly, cardiogenesis-specific transcriptional regulation of *LGR5* expression was studied. *In silico* analysis revealed a putative novel human-specific *MESP1* binding site upstream of the transcription start site of the *LGR5* promoter, as well as a *LEF1* binding site upstream of the transcription start site of the *ISL1* promoter. The *in silico* findings were validated through chromatin immunoprecipitation assays using chromatin derived from cells collected on days 0 and 3 of cardiac differentiation, which indicate recruitment of MESP1 on *LGR5* promoter and LEF1 on *ISL1* promoter. The assay was repeated using the *LGR5* knockout hESCs, which showed decreased LEF1 binding to the *ISL1* promoter, suggesting that LGR5 plays a role in LEF1-mediated activation of *ISL1* expression. Taken together, these results suggest that LGR5-signaling and transcriptional regulation mediated by ISL1 are interconnected and play a key role in the differentiation and development of the progenitors of cono-ventricular myocytes.

Comparison between the *in vitro* (bulk RNA-seq) and *in vivo* (single-cell) populations indicated similarity in the transcriptional profiles between the *in vitro*-derived multipotent cardiac progenitors/cardiac intermediates and *in vivo* cono-ventricular progenitors. Likewise, the transcriptional profiles of *in vitro* cardiomyocytes resembled those of the *in vivo* free-ventricular myocytes. To pinpoint *in vivo* cardiac populations (cono-ventricular vs free-wall ventricular) that may be differentially affected in CHD, a binomial test was carried out. Based on their expression of CHD-associated genes (identified by the PCGC), cono-ventricular progenitors and cono-ventricular myocytes appear to be a hotspot in the pathogenesis of CHD.

To conclude, LGR5 is a novel marker and regulator of the ISL1+ cardiac cono-ventricular myocyte lineage. Given the central role of myogenic progenitors of the outflow tract in the pathogenesis of conotruncal cardiac defects, this population may hold an important key to understanding and modeling CHD.

8 GENERAL DISCUSSION AND OUTLOOK

The studies constituting this doctoral thesis provide new insights into the development of cardiac progenitors and shed light on their role in the pathogenesis of complex CHD. Specifically, we advanced the understanding of cardiac progenitors by:

1. Building a differentiation roadmap of hESCs to cardiomyocytes, which revealed the developmental relationships and dynamics of the cardiomyocyte lineage and other cell types.
2. Defining the embryonic cardiac cell types, developmental timepoints and cardiac progenitor-related gene ensembles contributing to the development of non-syndromic Tetralogy of Fallot.
3. Identifying and characterizing a novel cardiac progenitor population, which might play a central role in the development and abnormalities of the human cardiac outflow tract.

A main premise of this thesis is that the hESC cardiac differentiation closely recapitulates the genetic programs of human embryonic cardiogenesis. In recent years, there has been a growing interest in studying early heart development, as well as the adult heart, at single-cell resolution (Paik et al. 2020). The heart progenitors, which are the focus of this thesis, are of particular interest due to their applicability in understanding human heart development, disease modeling and cell replacement therapy (Protze et al. 2019). Recent single-cell analyses of cardiac differentiation intermediates from hPSCs (Friedman et al. 2018; Churko et al. 2018; Ruan et al. 2019) and pre-natal human hearts (Asp et al. 2019; Cui et al. 2019; Suryawanshi et al. 2020), have revealed novel insights into the early cardiac progenitor populations as well as embryonic and fetal cardiac cell types. To advance the field, we used high temporal resolution to build a connected roadmap of the hESC cardiac differentiation (paper I) and characterized the cono-ventricular and free-wall ventricular populations in the human embryonic heart in great detail (paper III). Taken together, these studies provide both an extensive resource of transcriptional signatures and a detailed characterization of the gene expression dynamics of human cardiac lineage specification.

In paper I, we conducted a single-cell analysis of multiple differentiation timepoints, which allowed us to build a connected roadmap of the hESC cardiogenesis. It is evident from our results and other recent single-cell studies, that we should think of cardiac progenitors passing through momentary transitional stages rather than having a fixed transcriptional signature (Tyser et al. 2021). Therefore, pinpointing and defining a cardiac progenitor with a specific transcriptional signature, which can for example be robustly generated for cell replacement therapies, is a challenge. Although the field of cardiovascular regeneration has not yet reached a consensus about the optimal hPSC-derived cell type for cardiac cell replacement therapy (Chien et al. 2019), the essential first step is to fully understand the differentiation process. Together with previous studies, our study provides a guide for outlining specific progenitor stages, and visualizes the branching points and onset of specification along the

differentiation trajectory of the cardiac lineage. In particular, the RNA velocity and trajectory analysis in our study provides further information on the maintenance of an early progenitor-like identity (stemness) versus cardiac commitment, which could prove valuable as current research efforts are attempting to prolong the progenitor-like proliferative state to increase the graft size upon injection of hPSC-derived cardiac progenitors into infarcted hearts (Schwach et al. 2020).

Moreover, the study presented in paper I adds new insights into the hESC-derived non-myogenic populations generated alongside cardiomyocytes. Fibroblasts in particular are an integral component of the myocardium (Moore-Morris et al. 2015), and our study defines their identity in the hESC cardiac differentiation system. In terms of future applications, the ability to generate cardiac fibroblast-like cells *in vitro* facilitates research on cardiac development and disease, as well as provides cells for cardiac tissue engineering (Zhang et al. 2019). Additionally, our study suggests, that ISL1 transcription factor may play a role in the decision between cardiac and neural lineages. In order to gain a full picture of the segregation of the neural-like population in the *ISL1* knockout differentiation, single-cell RNA sequencing of cells collected on days 3, 4, and 5 of the *ISL1* knockout differentiation could be conducted. Both a previous study and our results indicate that *ISL1* expression is not necessary for cardiomyocyte differentiation (Quaranta et al. 2018). Our study therefore rather suggests that it contributes to the differentiation process as a partial molecular switch between lineages.

The second main premise of this thesis is that single-cell analysis of human cardiac progenitor lineages provides a framework for investigating the pathogenesis of CHD. Complex forms of CHD are rarely curable, and although a large body of research has addressed the pathogenesis of CHD, it remains largely elusive (Nees and Chung 2020). Understanding the pathophysiology through genetics and cell and molecular biology, along with thorough clinical characterization, is the first step towards improved personalized interventions. Next-generation sequencing has opened new possibilities for mapping the mutational landscapes of various types of structural CHD (Shabana et al. 2020). Our knowledge on the Mendelian, *de novo* and rare recessive variants causing CHD and Tetralogy of Fallot specifically is rapidly growing due to the recent and ongoing research efforts (paper II, Jin et al. 2017; Page et al. 2019; Richter et al. 2020). The diversity of genetic mutations found suggests that the genetic landscape of CHD is highly complex and context-dependent, and there is still much to be discovered. In particular, the role of mutations within the non-coding genome in CHD pathogenesis is elusive (Richter et al. 2020), and we plan to explore these in a future study based on the whole-genome sequencing analysis of the Tetralogy of Fallot cohort.

Cardiac outflow tract development is the focal point underlying the pathogenesis of many types of complex CHD (Neeb et al. 2103; Nakajima 2010). Both defective alignment and incomplete septation can lead to abnormal outflow tract development, and previous studies have associated these disease processes with the SHF and neural crest progenitors (Neeb et al.

2013). To shed light on the connection between conotruncal disease phenotypes (Tetralogy of Fallot in particular) and early cardiac populations, we intersected CHD genetics with single-cell analyses of human heart development *in vivo* and *in vitro*, as presented in papers II and III. Specifically, we found that the myogenic cells building the cardiac chambers, and notably the outflow tract, are mainly affected by the Tetralogy of Fallot-associated genetic mutations. No enrichment of Tetralogy of Fallot-associated genes was found among the non-myogenic cell types, such as neural crest cells, highlighting the importance of the myogenic cardiac progenitors and cardiomyocytes in the pathogenesis of non-syndromic cases of Tetralogy of Fallot.

The LGR5+ cono-ventricular progenitors emerged from these analyses as a novel subset of ISL1+ SHF derivatives, which likely play a central role in the outflow tract morphogenesis and pathogenesis of conotruncal cardiac malformations. The cono-ventricular progenitors contribute to the outflow tract myocardium and are distinct from the progenitors that form the right ventricle (Nakajima 2010), i.e. the free-wall ventricular cardiomyocytes, as discussed in paper III. In addition to specifically marking the cono-ventricular progenitor population, LGR5 plays a role in their lineage progression, as deletion of *LGR5* resulted in blocking the molecular machinery promoting cardiac differentiation.

It would be interesting to study whether the LGR5+ cono-ventricular progenitor population is affected in the embryonic/fetal hearts with Tetralogy of Fallot phenotype. This is likely the case since they express several genes indicated in CHD and Tetralogy of Fallot specifically, as shown in papers II and III. Future studies are needed to understand the differential roles of LGR5+ and LGR5- progenitor-derived cardiac lineages in the heart development, in particular in the development of cono-ventricular versus free-wall ventricular cardiomyocytes. However, the LGR5+ cardiac population is not present in the developing mouse heart, and hence cannot be studied in mouse models by lineage tracing. Instead, hESC-cardiac differentiation and non-human primate cardiogenesis could be used as model systems to further study the involvement of LGR5+ progenitors in the CHD pathogenesis.

Establishing the onset and timeline of CHD in the early heart development, especially in human, is one of the fundamental steps towards understanding the CHD pathogenesis. In paper II, we used the hESC differentiation roadmap presented in paper I to pinpoint developmental (temporal) timepoints along the pluripotency-to-cardiomyocyte transition affected by Tetralogy of Fallot-associated genetic mutations. Our approach was adapted from a previous study, which investigated the timing of neurological disease-associated genes based on single-cell RNA sequencing analysis of the developing human brain (Li et al. 2018). We found that subsets of the disease-associated genes were expressed at specific stages of the cardiomyocyte differentiation, both at the pluripotent/pre-cardiac stage and at the late cardiac progenitor/cardiomyocyte stage. A previous study, which intersected CHD-associated genes with a time-course bulk transcriptional analysis of cardiac development in mouse, arrived to similar conclusions (Li et al. 2014). While these findings indicate that the genetics and pathogenesis of Tetralogy of Fallot affect multiple stages of cardiac development, we were

able to identify specific subsets of disease-associated genes specifically expressed by the cardiac progenitors. The role of many of these genes in cardiac development is currently unknown, and therefore they are interesting candidates for hPSC- and animal model-based loss- and gain-of-function studies.

Finally, to put our findings in a wider context, we compared the genetic landscape of Chinese Tetralogy of Fallot cases with a previously reported cohort of European descent (Jin et al. 2017). Although we observed an overlap in the genetic findings from these two cohorts, we also observed discrepancies, in particular related to the molecular and cellular functions of the genes indicated in the two patient populations. Previous studies have indicated that ethnicity contributes to the differences between patient phenotypes and genetics in complex CHD (Van Der Linde et al. 2011; Freeman et al. 2008). In particular, the occurrence of left versus right sided heart phenotypes differs between patients with complex CHD of Asian and European descent (Zimmerman et al. 2017; Rokicki et al. 2003). Our findings support these notions regarding the genetic landscapes of Tetralogy of Fallot.

In the research presented in this thesis, we have applied *in vitro* cardiogenesis using hESCs as a platform for understanding early lineage segregation and developmental progenitor dynamics. Furthermore, we have built comprehensive transcriptional datasets of the early human *in vitro* and *in vivo* cardiac populations. Together with genetic analysis of complex CHD, this data highlights the role of cardiac progenitors in the pathogenesis of CHD. The currently ongoing research efforts to further characterize the early human and non-human primate heart development will continue to advance our understanding of cardiogenesis, and might eventually solve the riddle around CHD.

9 ACKNOWLEDGEMENTS

I would like to thank my supervisor, **Professor Kenneth Chien**, for giving me the opportunity to join the group and work on these interesting projects. Thanks to your guidance and encouragement I have become the independent researcher I am today.

Thanks to my co-supervisors **Christian Riedel** and **Fredrik Lanner** for your support, help and inspiration.

As a proud member of **the Chien lab** family I would like to thank all present and past colleagues for making my time in the lab a unique experience.

Chuen Yan, thank you for being there for me through my PhD journey. I literally don't know how it would have turned out otherwise. **Elif**, thank you for being a supportive and caring colleague. You are a source of inspiration! Special thanks for reviewing my thesis. **Raffa**, thank you for the lovely chats and really being a great friend at work. **Xiaobing**, you are an irreplaceable part of the Chien lab. You keep the lab organized and create a positive atmosphere in the office. **Alex**, thank you for thoroughly reviewing my thesis, for friendly chats and bringing pad thai to hungry people! It has been fun to work together. **Katarina**, thank you for being a helpful and caring colleague. Many thanks for all the help with the ethical application, paper publication and all the other things. **Jesper**, from cat facts to ethical applications, I know I can count on you. I wish all the best for you and Christina! **Makoto**, thank you for helpful comments on my manuscript, help with experiments and entertaining lab meetings. **Kylie**, thank you for being a great colleague and saving the day multiple times by bringing some Swedish fika to the office. **Nevin**, thank you for creating the unique and fun atmosphere of the Chien lab. **Yao**, thank you for being a great colleague and helping me with the cell culture work. Cell culture help may sound like a little thing, but it is not when the cells need one's attention every single day, weekends included. **Rui**, I have enjoyed sharing the cell culture room with you and our little chats in between medium changes. **Ran**, thank you for being a nice colleague and special thanks for our little boat trip to Finland. I am happy you appreciate the traditional Finnish fish soup! **Tamara**, thank you for all the nice conversations (and also for bringing Austrian snacks to the office). **Chimezie**, my office desk neighbor, thank you for sharing your positive attitude as well as all the delicious homemade cakes. **JP**, it has been nice to know you for this short period of time. I am sure you will make a great addition to the Chien lab.

Thanks to my fellow PhD student of the Chien lab. **Sunny**, thank you for introducing me to the Chien lab and your guidance during my master's studies. **Chikai**, thank you for being such a positive and helpful colleague. I wish you luck with your soon-to-be PhD defense. I'm sure you will do great! **Edu**, thanks for encouraging me to join the Chien lab. **Chris**, thank you for interesting lunch conversations, and special thanks for creating the office jungle.

Previous colleagues **Gianluigi**, **Jon**, **Wei**, **Jay** and **Jiejia**, thank you for bringing your own unique contributions to the Chien lab. **Kristine**, thank you for being such a lovely colleague

at work and organizing epic lab parties in your beautiful home. Me and others have missed you a lot. **Miia**, my Finnish colleague, I wish our time in the lab would have overlapped for a longer time. **Fede**, I have missed you in the lab. Thank you for always being so positive and friendly, and thanks for the lovely company in Heidelberg. **Erwin**, thank you for great chats and company during long evenings and weekends in the lab. When the -80 freezers broke down, we made a pretty efficient team to solve the problem. **Hasina** and **Kalai**, thank you for technical assistance and help in the lab.

Matti Nikkola, **Linda Lindell** and **Lina Petterson**, thank you for helping me to navigate through my PhD journey.

Dear collaborators, **Peter Gruber**, **Paul Tam**, **Merce Garcia-Barceló**, **Clara Tang**, **Mandy Lam**, **Sunny Zhuang**, **Haifa Hong** and **Peter Jin**, it has been fantastic working with you. I have learned a great deal about the clinical and genetic aspects of CHD from you.

Colleagues at **Integrated Cardio Metabolic Centre**, thanks to all of you. **Christina S** and **Angelina**, I am so happy we met and could share our PhD journeys. A lot has happened in our lives during this time, and it has been great to have such supportive and compassionate friends by my side. **Åsa**, **Elisabeth** and **Malin**, thank you for all the everyday support and for making ICMC such a lovely workplace. I would like to thank **Gizem**, **Christina J**, **Nelly**, **Aldo**, **Xin**, **Emmanuelle**, **Connie**, **Emelie**, **Ilke**, **Aldo**, **Bora**, **Xidan**, **Katja**, **André**, **Stefanos**, **Jérôme** and the whole ICMC crew for making the fikas, lunches, department pubs and parties (and also work) so much fun.

Thanks to **Leena** and **Catalin** for taking me to nature adventures around Stockholm. More than that, thanks for being great friends and brightened up the last year of my PhD (and my experience of the global pandemic)!

Iida Valkonen, thank you for the cover illustration.

Finally, I am so appreciative to have the supportive and caring family that I do.

Kiitos **äiti**. Olet aina ollut tukena kaikissa edesottamuksissani. On ollut ihana aina puhua puhelimessa ja kertoa sinulle huolista, kun vastauksesi on, että kyllä sinä pärjäät, luotan sinuun, älä huoli. Ja niin vaan olen pärjännyt sinun kannustamana ♥. Kiitos **isä**. Olet kannustanut minua aina seuraamaan unelmiani. Sain sinulta aikoinaan idean opiskella Karoliinisessa instituutissa ja sinun uurastamisesi väitöskirjasi parissa inspiroi osaltaan minutkin jaksamaan tämän urakan loppuun asti. Ehkä se purjevenekin joskus tulee hankittua. Rakkaat siskot **Sara** ja **Rita**, kiitos, aina kun nähdään, niin tuntuu kuin oltaisi nähty viimeksi eilen. **Tiina**, kiitos, että olen inspiroinut minua lukemaan muutakin kuin tiedejulkaisuja.

10 REFERENCES

Aggarwal, V. S., & Morrow, B. E. (2008). Genetic modifiers of the physical malformations in velo-cardio-facial syndrome/DiGeorge syndrome. *Developmental Disabilities Research Reviews*, *14*(1), 19–25. https://doi.org/10.1002/ddrr.4

Ai, D., Liu, W., Ma, L., Dong, F., Lu, M.-F., Wang, D., Verzi, M. P., Cai, C., Gage, P. J., Evans, S., Black, B. L., Brown, N. A., & Martin, J. F. (2006). Pitx2 regulates cardiac left-right asymmetry by patterning second cardiac lineage-derived myocardium. *Developmental Biology*, *296*(2), 437–449. https://doi.org/10.1016/j.ydbio.2006.06.009

Aibar, S., González-Blas, C. B., Moerman, T., Huynh-Thu, V. A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J. C., Geurts, P., Aerts, J., Van Den Oord, J., Atak, Z. K., Wouters, J., & Aerts, S. (2017). SCENIC: Single-cell regulatory network inference and clustering. *Nature Methods*, *14*(11), 1083–1086. https://doi.org/10.1038/nmeth.4463

Akhirome, E., Walton, N. A., Nogee, J. M., & Jay, P. Y. (2017). The Complex Genetic Basis of Congenital Heart Defects. *Circulation Journal : Official Journal of the Japanese Circulation Society*, *81*(5), 629–634. https://doi.org/10.1253/circj.CJ-16-1343

Ali, S. R., Ranjbarvaziri, S., Talkhabi, M., Zhao, P., Subat, A., Hojjat, A., Kamran, P., Müller, A. M. S., Volz, K. S., Tang, Z., Red-Horse, K., & Ardehali, R. (2014). Developmental heterogeneity of cardiac fibroblasts does not predict pathological proliferation and activation. *Circulation Research*, *115*(7), 625–635. https://doi.org/10.1161/CIRCRESAHA.115.303794

Anderson, R. H., Mori, S., Spicer, D. E., Brown, N. A., & Mohun, T. J. (2016). Development and Morphology of the Ventricular Outflow Tracts. In *World journal for pediatric & congenital heart surgery* (Vol. 7, Issue 5, pp. 561–577). World J Pediatr Congenit Heart Surg. https://doi.org/10.1177/2150135116651114

Anderson, R. H., Spicer, D. E., Brown, N. A., & Mohun, T. J. (2014). The development of septation in the four-chambered heart. *Anatomical Record*, *297*(8), 1414–1429. https://doi.org/10.1002/ar.22949

Anderson, R. H., & Weinberg, P. M. (2005). The clinical anatomy of tetralogy of Fallot. In *Cardiology in the Young* (Vol. 15, Issue SUPPL. 1, pp. 38–47). Cardiol Young. https://doi.org/10.1017/S1047951105001010

Andrews, T. S., & Hemberg, M. (2018). Identifying cell populations with scRNASeq. In *Molecular Aspects of Medicine* (Vol. 59, pp. 114–122). Elsevier Ltd. https://doi.org/10.1016/j.mam.2017.07.002

Apitz, C., Webb, G. D., & Redington, A. N. (2009). Tetralogy of Fallot. In *The Lancet* (Vol. 374, Issue 9699, pp. 1462–1471). Lancet. https://doi.org/10.1016/S0140-6736(09)60657-7

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. In *Nature Genetics* (Vol. 25, Issue 1, pp. 25–29). Nature Publishing Group. https://doi.org/10.1038/75556

Ashraf, H., Pradhan, L., Chang, E. I., Terada, R., Ryan, N. J., Briggs, L. E., Chowdhury, R., Zárate, M. A., Sugi, Y., Nam, H. J., Benson, D. W., Anderson, R. H., & Kasahara, H. (2014). A mouse model of human congenital heart disease high incidence of diverse cardiac anomalies and ventricular noncompaction produced by heterozygous Nkx2-5 homeodomain missense mutation. *Circulation: Cardiovascular Genetics*, *7*(4), 423–433. https://doi.org/10.1161/CIRCGENETICS.113.000281

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., Wärdell, E., Custodio, J., Reimegård, J., Salmén, F., Österholm, C., Ståhl, P. L., Sundström, E., Åkesson, E., Bergmann, O., Bienko, M., Månsson-Broberg, A., Nilsson, M., Sylvén, C., & Lundeberg, J. (2019). A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. *Cell*, *179*(7), 1647-1660.e19. https://doi.org/10.1016/j.cell.2019.11.025

Azhar, M., Schultz, J. E. J., Grupp, I., Dorn, G. W., Meneton, P., Molin, D. G. M., Gittenberger-de Groot, A. C., & Doetschman, T. (2003). Transforming growth factor beta in cardiovascular development and function. In *Cytokine and Growth Factor Reviews* (Vol. 14, Issue 5, pp. 391–407). Elsevier BV. https://doi.org/10.1016/S1359-6101(03)00044-3

Azhar, M., Wang, P.-Y., Frugier, T., Koishi, K., Deng, C., Noakes, P. G., & McLennan, I. S. (2010). Myocardial deletion of Smad4 using a novel α skeletal muscle actin Cre recombinase transgenic mouse causes misalignment of the cardiac outflow tract. *International Journal of Biological Sciences*, *6*(6), 546–555. https://doi.org/10.7150/ijbs.6.546

Bageritz, J., & Raddi, G. (2019). Single-cell RNA sequencing with drop-seq. In *Methods in Molecular Biology* (Vol. 1979, pp. 73–85). Humana Press Inc. https://doi.org/10.1007/978-1-4939-9240-9_6

Bailliard, F., & Anderson, R. H. (2009). Tetralogy of Fallot. *Orphanet Journal of Rare Diseases*, *4*(1). https://doi.org/10.1186/1750-1172-4-2

Bamforth, S. D., Bragança, J., Eloranta, J. J., Murdoch, J. N., Marques, F. I., Kranc, K. R., Farza, H., Henderson, D. J., Hurst, H. C., & Bhattacharya, S. (2001). Cardiac malformations, adrenal agenesis, neural crest defects and exencephaly in mice lacking Cited2, a new Tfap2 co-activator. *Nature Genetics*, *29*(4), 469–474. https://doi.org/10.1038/ng768

Banerjee, I., Fuseler, J. W., Price, R. L., Borg, T. K., & Baudino, T. A. (2007). Determination of cell types and numbers during cardiac development in the neonatal and adult rat and mouse. *American Journal of Physiology - Heart and Circulatory Physiology*, *293*(3). https://doi.org/10.1152/ajpheart.00514.2007

Barker, N., & Clevers, H. (2010). Leucine-Rich Repeat-Containing G-Protein-Coupled Receptors as Markers of Adult Stem Cells. In *Gastroenterology* (Vol. 138, Issue 5, pp. 1681–1696). W.B. Saunders. https://doi.org/10.1053/j.gastro.2010.03.002

Barnes, R. M., Firulli, B. A., Conway, S. J., Vincentz, J. W., & Firulli, A. B. (2010). Analysis of the Hand1 cell lineage reveals novel contributions to cardiovascular, neural crest, extra-embryonic, and lateral mesoderm derivatives. *Developmental Dynamics*, *239*(11), 3086–3097. https://doi.org/10.1002/dvdy.22428

Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W. H., Ng, L. G., Ginhoux, F., & Newell, E. W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*, *37*(1), 38–47. https://doi.org/10.1038/nbt.4314

Behrens, A. N., Iacovino, M., Lohr, J. L., Ren, Y., Zierold, C., Harvey, R. P., Kyba, M., Garry, D. J., & Martin, C. M. (2013). Nkx2-5 mediates differential cardiac differentiation through interaction with Hoxa10. *Stem Cells and Development*, *22*(15), 2211–2220. https://doi.org/10.1089/scd.2012.0611

Ben Jehuda, R., Shemer, Y., & Binah, O. (2018). Genome Editing in Induced Pluripotent Stem Cells using CRISPR/Cas9. In *Stem Cell Reviews and Reports* (Vol. 14, Issue 3, pp. 323–336). Humana Press Inc. https://doi.org/10.1007/s12015-018-9811-3

Benson, D. W., Silberbach, G. M., Kavanaugh-McHugh, A., Cottrill, C., Zhang, Y., Riggs, S., Smalls, O., Johnson, M. C., Watson, M. S., Seidman, J. G., Seidman, C. E., Plowden, J., & Kugler, J. D. (1999). Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways. *Journal of Clinical Investigation*, *104*(11), 1567–1573. https://doi.org/10.1172/JCI8154

Bergwerff, M., Verberne, M. E., DeRuiter, M. C., Poelmann, R. E., & Gittenberger-de Groot, A. C. (1998). Neural crest cell contribution to the developing circulatory system: implications for vascular morphology? *Circulation Research*, *82*(2), 221–231. https://doi.org/10.1161/01.res.82.2.221

Bertrand, N., Roux, M., Ryckebüsch, L., Niederreither, K., Dollé, P., Moon, A., Capecchi, M., & Zaffran, S. (2011). Hox genes define distinct progenitor sub-domains within the second heart field. *Developmental Biology*, *353*(2), 266–274. https://doi.org/10.1016/j.ydbio.2011.02.029

Biben, C., & Harvey, R. P. (1997). Homeodomain factor Nkx2-5 controls left/right asymmetric expression of bHLH gene eHAND during murine heart development. *Genes and Development*, *11*(11), 1357–1369. https://doi.org/10.1101/gad.11.11.1357

Blakeley, P., Fogarty, N. M. E., Del Valle, I., Wamaitha, S. E., Hu, T. X., Elder, K., Snell, P., Christie, L., Robson, P., & Niakan, K. K. (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development (Cambridge)*, *142*(18), 3151–3165. https://doi.org/10.1242/dev.123547

Boot, M. J., Gittenberger-De Groot, A. C., Van Iperen, L., Hierck, B. P., & Poelmann, R. E. (2003). Spatiotemporally Separated Cardiac Neural Crest Subpopulations That Target the Outflow Tract Septum and Pharyngeal Arch Arteries. *Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology*, *275*(1), 1009–1018. https://doi.org/10.1002/ar.a.10099

Boroviak, T., Stirparo, G. G., Dietmann, S., Hernando-Herraez, I., Mohammed, H., Reik, W., Smith, A., Sasaki, E., Nichols, J., & Bertone, P. (2018). Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development (Cambridge)*, *145*(21). https://doi.org/10.1242/dev.167833

Brady, G., Barbara, M., & Iscove, N. N. (1990). *Representative in Vitro cDNA Amplification From Individual Hemopoietic Cells and Colonies*.

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csárdi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A., Kircher, M., Albert, F. W., Zeller, U., Khaitovich, P., Grützner, F., Bergmann, S., Nielsen, R., Pääbo, S., & Kaessmann, H. (2011). The evolution of gene expression levels in mammalian organs. *Nature*, *478*(7369), 343–348. https://doi.org/10.1038/nature10532

Brennan, J., Lu, C. C., Norris, D. P., Rodriguez, T. A., Beddington, R. S. P., & Robertson, E. J. (2001). Nodal signalling in the epiblast patterns the early mouse embryo. *Nature*, *411*(6840), 965–969. https://doi.org/10.1038/35082103

Brewer, S., Jiang, X., Donaldson, S., Williams, T., & Sucov, H. M. (2002). Requirement for AP-2alpha in cardiac outflow tract morphogenesis. *Mechanisms of Development*, *110*(1–2), 139–149. https://doi.org/10.1016/s0925-4773(01)00579-2

Brown, C. B., Boyer, A. S., Runyan, R. B., & Barnett, J. V. (1996). Antibodies to the type II TGFβ receptor block cell activation and migration during atrioventricular cushion transformation in the heart. *Developmental Biology*, *174*(2), 248–257. https://doi.org/10.1006/dbio.1996.0070

Bruneau, B. G. (2013). Signaling and transcriptional networks in heart development and regeneration. *Cold Spring Harbor Perspectives in Biology*, *5*(3), a008292. https://doi.org/10.1101/cshperspect.a008292

Bruneau, B. G., Nemer, G., Schmitt, J. P., Charron, F., Robitaille, L., Caron, S., Conner, D. A., Gessler, M., Nemer, M., Seidman, C. E., & Seidman, J. G. (2001). A murine model of Holt-Oram syndrome defines roles of the T-Box transcription factor Tbx5 in cardiogenesis and disease. *Cell*, *106*(6), 709–721. https://doi.org/10.1016/S0092-8674(01)00493-7

Bu, L., Jiang, X., Martin-Puig, S., Caron, L., Zhu, S., Shao, Y., Roberts, D. J., Huang, P. L., Domian, I. J., & Chien, K. R. (2009). Human ISL1 heart progenitors generate diverse multipotent cardiovascular cell lineages. *Nature*, *460*(7251), 113–117. https://doi.org/10.1038/nature08191

Buckingham, M., Meilhac, S., & Zaffran, S. (2005). Building the mammalian heart from two sources of myocardial cells. In *Nature Reviews Genetics* (Vol. 6, Issue 11, pp. 826–835). https://doi.org/10.1038/nrg1710

Buijtendijk, M. F. J., Barnett, P., & Hoff, M. J. B. (2020). Development of the human heart. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *184*(1), 7–22. https://doi.org/10.1002/ajmg.c.31778

Burridge, P. W., Thompson, S., Millrod, M. A., Weinberg, S., Yuan, X., Peters, A., Mahairaki, V., Koliatsos, V. E., Tung, L., & Zambidis, E. T. (2011). A universal system for highly efficient cardiac differentiation of human induced pluripotent stem cells that eliminates interline variability. *PloS One*, *6*(4), e18293. https://doi.org/10.1371/journal.pone.0018293

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, *36*(5), 411–420. https://doi.org/10.1038/nbt.4096

Cai, C. L., Liang, X., Shi, Y., Chu, P. H., Pfaff, S. L., Chen, J., & Evans, S. (2003). Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Developmental Cell*, *5*(6), 877–889. https://doi.org/10.1016/S1534-5807(03)00363-0

Calcagni, G., Unolt, M., Digilio, M. C., Baban, A., Versacci, P., Tartaglia, M., Baldini, A., & Marino, B. (2017). Congenital heart disease and genetic syndromes: new insights into molecular mechanisms. *Expert Review of Molecular Diagnostics*, *17*(9), 861–870. https://doi.org/10.1080/14737159.2017.1360766

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., Trapnell, C., & Shendure, J. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, *566*(7745), 496–502. https://doi.org/10.1038/s41586-019-0969-x

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., Trapnell, C., & Shendure, J. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, *566*(7745), 496–502. https://doi.org/10.1038/s41586-019-0969-x

Carmeliet, P. (2000). Mechanisms of angiogenesis and arteriogenesis. In *Nature Medicine* (Vol. 6, Issue 4, pp. 389–395). Nature Publishing Group. https://doi.org/10.1038/74651

Chahal, G., Tyagi, S., & Ramialison, M. (2019). Navigating the non-coding genome in heart development and Congenital Heart Disease. In *Differentiation* (Vol. 107, pp. 11–23). Elsevier Ltd. https://doi.org/10.1016/j.diff.2019.05.001

Chen, G., Ning, B., & Shi, T. (2019). Single-cell RNA-seq technologies and related computational data analysis. In *Frontiers in Genetics* (Vol. 10, Issue APR). Frontiers Media S.A. https://doi.org/10.3389/fgene.2019.00317

Chien, K. R., Frisén, J., Fritsche-Danielson, R., Melton, D. A., Murry, C. E., & Weissman, I. L. (2019). Regenerating the field of cardiovascular cell therapy. *Nature Biotechnology*, *37*(3), 232–237. https://doi.org/10.1038/s41587-019-0042-1

Christoffels, V. M., Habets, P. E. M. H., Franco, D., Campione, M., De Jong, F., Lamers, W. H., Bao, Z. Z., Palmer, S., Biben, C., Harvey, R. P., & Moorman, A. F. M. (2000). Chamber formation and morphogenesis in the developing mammalian heart. *Developmental Biology*, *223*(2), 266–278. https://doi.org/10.1006/dbio.2000.9753

Churko, J. M., Garg, P., Treutlein, B., Venkatasubramanian, M., Wu, H., Lee, J., Wessells, Q. N., Chen, S. Y., Chen, W. Y., Chetal, K., Mantalas, G., Neff, N., Jabart, E., Sharma, A., Nolan, G. P., Salomonis, N., & Wu, J. C. (2018). Defining human cardiac transcription factor hierarchies using integrated single-cell heterogeneity analysis. *Nature Communications*, *9*(1), 4906. https://doi.org/10.1038/s41467-018-07333-4

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), 819–823. https://doi.org/10.1126/science.1231143

Cowan, J. R., & Ware, S. M. (2015). Genetics and Genetic Testing in Congenital Heart Disease. In *Clinics in Perinatology* (Vol. 42, Issue 2, pp. 373–393). W.B. Saunders. https://doi.org/10.1016/j.clp.2015.02.009

Creazzo, T. L., Godt, R. E., Leatherbury, L., Conway, S. J., & Kirby, M. L. (1998). Role of cardiac neural crest cells in cardiovascular development. *Annual Review of Physiology*, *60*, 267–286. https://doi.org/10.1146/annurev.physiol.60.1.267

Cui, Y., Zheng, Y., Liu, X., Yan, L., Fan, X., Yong, J., Hu, Y., Dong, J., Li, Q., Wu, X., Gao, S., Li, J., Wen, L., Qiao, J., & Tang, F. (2019). Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. *Cell Reports*, *26*(7), 1934-1950.e5. https://doi.org/10.1016/j.celrep.2019.01.079

D'Amour, K. A., Agulnick, A. D., Eliazer, S., Kelly, O. G., Kroon, E., & Baetge, E. E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature Biotechnology*, *23*(12), 1534–1541. https://doi.org/10.1038/nbt1163

De La Pompa, J. L., Timmerman, L. A., Takimoto, H., Yoshida, H., Elia, A. J., Samper, E., Potter, J., Wakeham, A., Marengere, L., Langille, B. L., Crabtree, G. R., & Mak, T. W. (1998). Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. *Nature*, *392*(6672), 182–186. https://doi.org/10.1038/32419

De Lange, F. J., Moorman, A. F. M., Anderson, R. H., Männer, J., Soufan, A. T., De Gier-De Vries, C., Schneider, M. D., Webb, S., Van Den Hoff, M. J. B., & Christoffels, V. M. (2004). Lineage and morphogenetic analysis of the cardiac valves. *Circulation Research*, *95*(6), 645–654. https://doi.org/10.1161/01.RES.0000141429.13560.cb

de Soysa, T. Y., Ranade, S. S., Okawa, S., Ravichandran, S., Huang, Y., Salunga, H. T., Schricker, A., del Sol, A., Gifford, C. A., & Srivastava, D. (2019). Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. In *Nature* (Vol. 572, Issue 7767, pp. 120–124). Nature Publishing Group. https://doi.org/10.1038/s41586-019-1414-x

DeLaughter, D. M., Bick, A. G., Wakimoto, H., McKean, D., Gorham, J. M., Kathiriya, I. S., Hinson, J. T., Homsy, J., Gray, J., Pu, W., Bruneau, B. G., Seidman, J. G., & Seidman, C. E. (2016). Single-Cell Resolution of Temporal Gene Expression during Heart Development. *Developmental Cell*, *39*(4), 480–490. https://doi.org/10.1016/j.devcel.2016.10.001

Dettman, R. W., Denetclaw, W. J., Ordahl, C. P., & Bristow, J. (1998). Common epicardial origin of coronary vascular smooth muscle, perivascular fibroblasts, and intermyocardial fibroblasts in the avian heart. *Developmental Biology*, *193*(2), 169–181. https://doi.org/10.1006/dbio.1997.8801

Digilio, M. C., Marino, B., Grazioli, S., Agostino, D., Giannotti, A., & Dallapiccola, B. (1996). Comparison of occurrence of genetic syndromes in ventricular septal defect with pulmonic stenosis (classic tetralogy of Fallot) versus ventricular septal defect with pulmonic atresia. *American Journal of Cardiology*, *77*(15), 1375–1376. https://doi.org/10.1016/S0002-9149(96)00212-3

Dixit, R., Narasimhan, C., Balekundri, V. I., Agrawal, D., Kumar, A., & Mohapatra, B. (2018). Functionally significant, novel GATA4 variants are frequently associated with Tetralogy of Fallot. *Human Mutation*, *39*(12), 1957–1972. https://doi.org/10.1002/humu.23620

Donovan, J., Kordylewska, A., Jan, Y. N., & Utset, M. F. (2002). Tetralogy of fallot and other congenital heart defects in Hey2 mutant mice. *Current Biology*, *12*(18), 1605–1610. https://doi.org/10.1016/S0960-9822(02)01149-1

Dor, Y. (2001). A novel role for VEGF in endocardial cushion formation and its potential contribution to congenital heart defects. *Development*, 128(9):1531-8

Drawnel, F. M., Boccardo, S., Prummer, M., Delobel, F., Graff, A., Weber, M., Gérard, R., Badi, L., Kam-Thong, T., Bu, L., Jiang, X., Hoflack, J. C., Kiialainen, A., Jeworutzki, E., Aoyama, N., Carlson, C., Burcin, M., Gromo, G., Boehringer, M., … Iacone, R. (2014). Disease modeling and phenotypic drug screening for diabetic cardiomyopathy using human induced pluripotent stem cells. *Cell Reports*, *9*(3), 810–820. https://doi.org/10.1016/j.celrep.2014.09.055

Du, Q., de la Morena, M. T., & van Oers, N. S. C. (2020). The Genetics and Epigenetics of 22q11.2 Deletion Syndrome. In *Frontiers in Genetics* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/fgene.2019.01365

Dyer, L. A., & Kirby, M. L. (2009). The role of secondary heart field in cardiac development. In *Developmental Biology* (Vol. 336, Issue 2, pp. 137–144). Academic Press Inc. https://doi.org/10.1016/j.ydbio.2009.10.009

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., & Coleman, P. (1992). Analysis of gene expression in single live neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(7), 3010–3014. https://doi.org/10.1073/pnas.89.7.3010

Eisenberg, L. M., & Markwald, R. R. (1995). Molecular regulation of atrioventricular valvuloseptal morphogenesis. In *Circulation Research* (Vol. 77, Issue 1, pp. 1–6). Lippincott Williams and Wilkins. https://doi.org/10.1161/01.RES.77.1.1

Engleka, K. A., Manderfield, L. J., Brust, R. D., Li, L., Cohen, A., Dymecki, S. M., & Epstein, J. A. (2012). Islet1 derivatives in the heart are of both neural crest and second heart field origin. *Circulation Research*, *110*(7), 922–926. https://doi.org/10.1161/CIRCRESAHA.112.266510

Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, *292*(5819), 154–156. https://doi.org/10.1038/292154a0

Fahed, A. C., Gelb, B. D., Seidman, J. G., & Seidman, C. E. (2013). Genetics of congenital heart disease: The glass half empty. In *Circulation Research* (Vol. 112, Issue 4, pp. 707–720). Lippincott Williams and Wilkins. https://doi.org/10.1161/CIRCRESAHA.112.300853

Fallot, A. (1888). Contribution a l'anatomie pathologique de lamaladie bleue (cyanose cardiaque). Marseille Med J 25:77–93,138–158, 207–223, 270–286, 341–354, 403–420.

Farrell, M. J., Burch, J. L., Wallis, K., Rowley, L., Kumiski, D., Stadt, H., Godt, R. E., Creazzo, T. L., & Kirby, M. L. (2001). FGF-8 in the ventral pharynx alters development of myocardial calcium transients after neural crest ablation. *Journal of Clinical Investigation*, *107*(12), 1509–1517. https://doi.org/10.1172/JCI9317

Ferencz, C., Rubin, J. D., Mccarter, R. J., Brenner, J. I., Neill, C. A., Perry, L. W., Hepner, S. I., & Downing, J. W. (1985). Congenital heart disease: Prevalence at livebirth: The baltimore-washington infant study. *American Journal of Epidemiology*, *121*(1), 31–36. https://doi.org/10.1093/oxfordjournals.aje.a113979

Ferrara, N., Gerber, H. P., & LeCouter, J. (2003). The biology of VEGF and its receptors. In *Nature Medicine* (Vol. 9, Issue 6, pp. 669–676). Nat Med. https://doi.org/10.1038/nm0603-669

Foo, K. S., Lehtinen, M. L., Leung, C. Y., Lian, X., Xu, J., Keung, W., Geng, L., Kolstad, T. R. S., Thams, S., Wong, A. O.-T., Wong, N., Bylund, K., Zhou, C., He, X., Jin, S.-B., Clarke, J., Lendahl, U., Li, R. A., Louch, W. E., & Chien, K. R. (2018). Human ISL1+ Ventricular Progenitors Self-Assemble into an In Vivo Functional Heart Patch and Preserve Cardiac Function Post Infarction. *Molecular Therapy : The Journal of the American Society of Gene Therapy*, *26*(7), 1644–1659. https://doi.org/10.1016/j.ymthe.2018.02.012

Franco, D., Meilhac, S. M., Christoffels, V. M., Kispert, A., Buckingham, M., & Kelly, R. G. (2006). *Left and right ventricular contributions to the formation of the interventricular septum in the mouse heart*. https://doi.org/10.1016/j.ydbio.2006.02.045

Freeman, S. B., Bean, L. H., Allen, E. G., Tinker, S. W., Locke, A. E., Druschel, C., Hobbs, C. A., Romitti, P. A., Royle, M. H., Torfs, C. P., Dooley, K. J., & Sherman, S. L. (2008). Ethnicity, sex, and the incidence of congenital heart defects: A report from the National Down Syndrome Project. *Genetics in Medicine*, *10*(3), 173–180. https://doi.org/10.1097/GIM.0b013e3181634867

Friedman, C. E., Nguyen, Q., Lukowski, S. W., Helfer, A., Chiu, H. S., Miklas, J., Levy, S., Suo, S., Han, J. D. J., Osteil, P., Peng, G., Jing, N., Baillie, G. J., Senabouth, A., Christ, A. N., Bruxner, T. J., Murry, C. E., Wong, E. S., Ding, J., … Palpant, N. J. (2018). Single-Cell Transcriptomic Analysis of Cardiac Differentiation from Human PSCs Reveals HOPX-Dependent Cardiomyocyte Maturation. *Cell Stem Cell*, *23*(4), 586-598.e8. https://doi.org/10.1016/j.stem.2018.09.009

Garcia-Martinez, V., & Schoenwolf, G. C. (1993). Primitive-streak origin of the cardiovascular system in avian embryos. *Developmental Biology*, *159*(2), 706–719. https://doi.org/10.1006/dbio.1993.1276

Garg, V. (2006). Insights into the genetic basis of congenital heart disease. In *Cellular and Molecular Life Sciences* (Vol. 63, Issue 10, pp. 1141–1148). Springer. https://doi.org/10.1007/s00018-005-5532-2

Gessert, S., & Kühl, M. (2010). The multiple phases and faces of Wnt signaling during cardiac differentiation and development. In *Circulation Research* (Vol. 107, Issue 2, pp. 186–199). Circ Res. https://doi.org/10.1161/CIRCRESAHA.110.221531

Gifford, C. A., Ranade, S. S., Samarakoon, R., Salunga, H. T., Yvanka De Soysa, T., Huang, Y., Zhou, P., Elfenbein, A., Wyman, S. K., Bui, Y. K., Cordes Metzler, K. R., Ursell, P., Ivey, K. N., & Srivastava, D. (2019). Oligogenic inheritance of a human heart disease involving a genetic modifier. *Science*, *364*(6443), 865–870. https://doi.org/10.1126/science.aat5056

Gilbert, S. F. (2003). Developmental biology. In *Sinauer Associates*.

Goldmuntz, E., Geiger, E., & Benson, D. W. (2001). NKX2.5 mutations in patients with tetralogy of fallot. *Circulation*, *104*(21), 2565–2568. https://doi.org/10.1161/hc4601.098427

Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A., & Ruddle, F. H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(12 II), 7380–7384. https://doi.org/10.1073/pnas.77.12.7380

Gore, A., Li, Z., Fung, H. L., Young, J. E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M. A., Kiskinis, E., Lee, J. H., Loh, Y. H., Manos, P. D., Montserrat, N., Panopoulos, A. D., Ruiz, S., Wilbert, M. L., Yu, J., Kirkness, E. F., … Zhang, K. (2011). Somatic coding mutations in human induced pluripotent stem cells. *Nature*, *471*(7336), 63–67. https://doi.org/10.1038/nature09805

Greenway, S. C., Pereira, A. C., Lin, J. C., Depalma, S. R., Israel, S. J., Mesquita, S. M., Ergul, E., Conta, J. H., Korn, J. M., McCarroll, S. A., Gorham, J. M., Gabriel, S., Altshuler, D. M., De Lourdes Quintanilla-Dieck, M., Artunduaga, M. A., Eavey, R. D., Plenge, R. M., Shadick, N. A., Weinblatt, M. E., … Seidman, C. E. (2009). De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nature Genetics*, *41*(8), 931–935. https://doi.org/10.1038/ng.415

Grunert, M., Appelt, S., Schönhals, S., Mika, K., Cui, H., Cooper, A., Cyganek, L., Guan, K., & Sperling, S. R. (2020). Induced pluripotent stem cells of patients with Tetralogy of Fallot reveal transcriptional alterations in cardiomyocyte differentiation. *Scientific Reports*, *10*(1), 10921. https://doi.org/10.1038/s41598-020-67872-z

Guadix, J. A., Orlova, V. V., Giacomelli, E., Bellin, M., Ribeiro, M. C., Mummery, C. L., Pérez-Pomares, J. M., & Passier, R. (2017). Human Pluripotent Stem Cell Differentiation into Functional Epicardial Progenitor Cells. *Stem Cell Reports*, *9*(6), 1754–1764. https://doi.org/10.1016/j.stemcr.2017.10.023

Gupta, M. K., De Jesus, D. F., Kahraman, S., Valdez, I. A., Shamsi, F., Yi, L., Swensen, A. C., Tseng, Y.-H., Qian, W.-J., & Kulkarni, R. N. (2018). Insulin receptor-mediated signaling regulates pluripotency markers and lineage differentiation. *Molecular Metabolism*, *18*, 153–163. https://doi.org/10.1016/j.molmet.2018.09.003

Gupta, R. M., & Musunuru, K. (2014). Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. In *Journal of Clinical Investigation* (Vol. 124, Issue 10, pp. 4154–4161). American Society for Clinical Investigation. https://doi.org/10.1172/JCI72992

Gustafsson, E., Brakebusch, C., Hietanen, K., & Fässler, R. (2001). Tie-1-directed expression of Cre recombinase in endothelial cells of embryoid bodies and transgenic mice. *Journal of Cell Science*, *114*(Pt 4), 671–676.

Habets, P. E. M. H., Moorman, A. F. M., Clout, D. E. W., Van Roon, M. A., Lingbeek, M., Van Lohuizen, M., Campione, M., & Christoffels, V. M. (2002). Cooperative action of Tbx2 and Nkx2.5 inhibits ANF expression in the atrioventricular canal: Implications for cardiac chamber formation. *Genes and Development*, *16*(10), 1234–1246. https://doi.org/10.1101/gad.222902

Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F., & Theis, F. J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nature Methods*, *13*(10), 845–848. https://doi.org/10.1038/nmeth.3971

Harvey, R. P. (2002). Patterning the vertebrate heart. In *Nature Reviews Genetics* (Vol. 3, Issue 7, pp. 544–556). Nat Rev Genet. https://doi.org/10.1038/nrg843

Hashimoto, T., Kojima, K., Otaka, A., Takeda, Y. S., Tomita, N., & Tamada, Y. (2013). Quantitative evaluation of fibroblast migration on a silk fibroin surface and TGFBI gene expression. *Journal of Biomaterials Science. Polymer Edition*, *24*(2), 158–169. https://doi.org/10.1163/156856212X629025

Hasten, E., McDonald-McGinn, D. M., Crowley, T. B., Zackai, E., Emanuel, B. S., Morrow, B. E., & Racedo, S. E. (2018). Dysregulation of TBX1 dosage in the anterior heart field results in congenital heart disease resembling the 22q11.2 duplication syndrome. *Human Molecular Genetics*, *27*(11), 1847–1857. https://doi.org/10.1093/hmg/ddy078

Hill, A. J., McFaline-Figueroa, J. L., Starita, L. M., Gasperini, M. J., Matreyek, K. A., Packer, J., Jackson, D., Shendure, J., & Trapnell, C. (2018). On the design of CRISPR-based single-cell molecular screens. *Nature Methods*, *15*(4), 271–274. https://doi.org/10.1038/nmeth.4604

Hill, M. C., Kadow, Z. A., Li, L., Tran, T. T., Wythe, J. D., & Martin, J. F. (2019). A cellular atlas of Pitx2-dependent cardiac development. *Development (Cambridge, England)*, *146*(12). https://doi.org/10.1242/dev.180398

Hinton, R. B. (2013). Genetic and environmental factors contributing to cardiovascular malformation: A unified approach to risk. In *Journal of the American Heart Association* (Vol. 2, Issue 3). https://doi.org/10.1161/JAHA.113.000292

Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., Cost, G. J., Zhang, L., Santiago, Y., Miller, J. C., Zeitler, B., Cherone, J. M., Meng, X., Hinkley, S. J., Rebar, E. J., Gregory, P. D., Urnov, F. D., & Jaenisch, R. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nature Biotechnology*, *29*(8), 731–734. https://doi.org/10.1038/nbt.1927

Hoffman, J. I. (1995). Incidence of congenital heart disease: II. Prenatal incidence. *Pediatric Cardiology*, *16*(4), 155–165. https://doi.org/10.1007/BF00794186

Hoffman, J. I. E., & Kaplan, S. (2002). The incidence of congenital heart disease. In *Journal of the American College of Cardiology* (Vol. 39, Issue 12, pp. 1890–1900). https://doi.org/10.1016/S0735-1097(02)01886-7

Hofmann, J. J., Briot, A., Enciso, J., Zovein, A. C., Ren, S., Zhang, Z. W., Radtke, F., Simons, M., Wang, Y., & Luisa Iruela-Arispe, M. (2012). Endothelial deletion of murine Jag1 leads to valve calcification and congenital heart defects associated with Alagille syndrome. *Development (Cambridge)*, *139*(23), 4449–4460. https://doi.org/10.1242/dev.084871

Hu, B. Y., Weick, J. P., Yu, J., Ma, L. X., Zhang, X. Q., Thomson, J. A., & Zhang, S. C. (2010). Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(9), 4335–4340. https://doi.org/10.1073/pnas.0910012107

Hwang, B., Lee, J. H., & Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. In *Experimental and Molecular Medicine* (Vol. 50, Issue 8). Nature Publishing Group. https://doi.org/10.1038/s12276-018-0071-8

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., Soreq, H., & Benvenisty, N. (2000). Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Molecular Medicine (Cambridge, Mass.)*, *6*(2), 88–95. https://doi.org/10.1007/bf03401776

Ivey, M. J., & Tallquist, M. D. (2016). Defining the cardiac fibroblast. In *Circulation Journal* (Vol. 80, Issue 11, pp. 2269–2276). Japanese Circulation Society. https://doi.org/10.1253/circj.CJ-16-1003

Jain, R., Engleka, K. A., Rentschler, S. L., Manderfield, L. J., Li, L., Yuan, L., & Epstein, J. A. (2011). Cardiac neural crest orchestrates remodeling and functional maturation of mouse semilunar valves. *The Journal of Clinical Investigation*, *121*(1), 422–430. https://doi.org/10.1172/JCI44244

Jenkins, K. J., Correa, A., Feinstein, J. A., Botto, L., Britt, A. E., Daniels, S. R., Elixson, M., Warnes, C. A., Webb, C. L., & American Heart Association Council on Cardiovascular Disease in the Young. (2007). Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation*, *115*(23), 2995–3014. https://doi.org/10.1161/CIRCULATIONAHA.106.183216

Ji, Z., & Ji, H. (2016). TSCAN: Pseudo-time reconstruction and evaluation in single-cell RNA-seq analysis. *Nucleic Acids Research*, *44*(13), e117. https://doi.org/10.1093/nar/gkw430

Jia, G., Preussner, J., Chen, X., Guenther, S., Yuan, X., Yekelchyk, M., Kuenne, C., Looso, M., Zhou, Y., Teichmann, S., & Braun, T. (2018). Single cell RNA-seq and ATAC-seq analysis of cardiac progenitor cell transition states and lineage settlement. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-07307-6

Jiang, X. (2000). Fate of the mammalian cardiac neural crest. *Development*, 127(8):1607-16

Jiang, Y., Park, P., Hong, S. M., & Ban, K. (2018). Maturation of cardiomyocytes derived from human pluripotent stem cells: Current strategies and limitations. In *Molecules and Cells* (Vol. 41, Issue 7, pp. 613–621). Korean Society for Molecular and Cellular Biology. https://doi.org/10.14348/molcells.2018.0143

Jin, S. C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., Depalma, S. R., Zeng, X., Qi, H., Chang, W., Sierant, M. C., Hung, W. C., Haider, S., Zhang, J., Knight, J., Bjornson, R. D., Castaldi, C., Tikhonoa, I. R., Bilguvar, K., Mane, S. M., … Brueckner, M. (2017). Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. *Nature Genetics*, *49*(11), 1593–1601. https://doi.org/10.1038/ng.3970

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Källén, B., Mastroiacovo, P., & Robert, E. (1996). Major congenital malformations in down syndrome. *American Journal of Medical Genetics*, *65*(2), 160–166. https://doi.org/10.1002/(SICI)1096-8628(19961016)65:2<160::AID-AJMG16>3.0.CO;2-O

Kattman, S. J., Witty, A. D., Gagliardi, M., Dubois, N. C., Niapour, M., Hotta, A., Ellis, J., & Keller, G. (2011). Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell*, 8(2):228-40. doi: 10.1016/j.stem.2010.12.008

Kaufman, M. H., & Navaratnam, V. (1981). Early differentiation of the heart in mouse embryos. *Journal of Anatomy*, *133*(Pt 2), 235–246. http://www.ncbi.nlm.nih.gov/pubmed/7333952

Kawasaki, T., Kitsukawa, T., Bekku, Y., Matsuda, Y., Sanbo, M., Yagi, T., & Fujisawa, H. (1999). A requirement for neuropilin-1 in embryonic vessel formation. *Development (Cambridge, England)*, *126*(21), 4895–4902.

Kehat, I., Kenyagin-Karsenti, D., Snir, M., Segev, H., Amit, M., Gepstein, A., Livne, E., Binah, O., Itskovitz-Eldor, J., & Gepstein, L. (2001). Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *Journal of Clinical Investigation*, *108*(3), 407–414. https://doi.org/10.1172/jci12131

Kelly, R. G., Brown, N. A., & Buckingham, M. E. (2001). The Arterial Pole of the Mouse Heart Forms from Fgf10-Expressing Cells in Pharyngeal Mesoderm. *Developmental Cell*, *1*(3), 435–440. https://doi.org/10.1016/S1534-5807(01)00040-5

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M. J., Ji, H., Ehrlich, L. I. R., Yabuuchi, A., Takeuchi, A., Cunniff, K. C., Hongguang, H., McKinney-Freeman, S., Naveiras, O., Yoon, T. J., Irizarry, R. A., Jung, N., … Daley, G. Q. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature*, *467*(7313), 285–290. https://doi.org/10.1038/nature09342

Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(3), 1156–1160. https://doi.org/10.1073/pnas.93.3.1156

Kinder, S. J., Tsang, T. E., Quinlan, G. A., Hadjantonakis, A. K., Nagy, A., & Tam, P. P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development*, *126*(21), 4691 LP – 4701. http://dev.biologists.org/content/126/21/4691.abstract

King, M. C., & Wilson, A. C. (1975). Evolution at two levels in humans and chimpanzees. *Science*, *188*(4184), 107–116. https://doi.org/10.1126/science.1090005

Kirby, M. L., Gale, T. F., & Stewart, D. E. (1983). Neural crest cells contribute to normal aorticopulmonary septation. *Science*, *220*(4601), 1059–1061. https://doi.org/10.1126/science.6844926

Kokkinopoulos, I., Ishida, H., Saba, R., Ruchaya, P., Cabrera, C., Struebig, M., Barnes, M., Terry, A., Kaneko, M., Shintani, Y., Coppen, S., Shiratori, H., Ameen, T., Mein, C., Hamada, H., Suzuki, K., & Yashiro, K. (2015). Single-cell expression profiling reveals a dynamic state of cardiac precursor cells in the early mouse embryo. *PLoS ONE*, *10*(10). https://doi.org/10.1371/journal.pone.0140831

Kruithof, B. P. T., Duim, S. N., Moerkamp, A. T., & Goumans, M.-J. (2012). *TGFb and BMP signaling in cardiac cushion formation: Lessons from mice and chicken*. https://doi.org/10.1016/j.diff.2012.04.003

Kurosawa, H., Morita, K., Yamagishi, M., Shimizu, S., Becker, A. E., Anderson, R. H., Bove, E. L., Lofland, G. K., & Brown, J. W. (1998). Conotruncal repair for tetralogy of Fallot: Midterm results. *Journal of Thoracic and Cardiovascular Surgery*, *115*(2), 351–360. https://doi.org/10.1016/S0022-5223(98)70279-X

Kurosawa, H., Imai, Y., Nakazawa, M., Momma, K., & Takao, A. (1988). Conotruncal Repair of Tetralogy of Fallot. *Annals of Thoracic Surgery*, *45*(6), 661–666. https://doi.org/10.1016/S0003-4975(10)64771-6

Kwon, C., Arnold, J., Hsiao, E. C., Taketo, M. M., Conklin, B. R., & Srivastava, D. (2007). Canonical Wnt signaling is a positive regulator of mammalian cardiac progenitors. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(26), 10894–10899. https://doi.org/10.1073/pnas.0704044104

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco, G., … Kharchenko, P. V. (2018). RNA velocity of single cells. In *Nature* (Vol. 560, Issue 7719, pp. 494–498). Nature Publishing Group. https://doi.org/10.1038/s41586-018-0414-6

Laksman, Z., Wauchop, M., Lin, E., Protze, S., Lee, J., Yang, W., Izaddoustdar, F., Shafaattalab, S., Gepstein, L., Tibbits, G. F., Keller, G., & Backx, P. H. (2017). Modeling Atrial Fibrillation using Human Embryonic Stem Cell-Derived Atrial Tissue. *Scientific Reports*, *7*(1), 5268. https://doi.org/10.1038/s41598-017-05652-y

Lalani, S. R., & Belmont, J. W. (2014). Genetic basis of congenital cardiovascular malformations. *European Journal of Medical Genetics*, *57*(8), 402–413. https://doi.org/10.1016/j.ejmg.2014.04.010

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*. https://doi.org/10.1186/1471-2105-9-559

Laugwitz, K.-L., Moretti, A., Lam, J., Gruber, P., Chen, Y., Woodard, S., Lin, L.-Z., Cai, C.-L., Lu, M. M., Reth, M., Platoshyn, O., Yuan, J. X.-J., Evans, S., & Chien, K. R. (2005). Postnatal isl1+ cardioblasts enter fully differentiated cardiomyocyte lineages. *Nature*, *433*(7026), 647–653. https://doi.org/10.1038/nature03215

Laurent, L. C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J. V., Lee, S., Barrero, M. J., Ku, S., Martynova, M., Semechkin, R., Galat, V., Gottesfeld, J., Belmonte, J. C. I., Murry, C., Keirstead, H. S., Park, H. S., … Loring, J. F. (2011). Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell*, *8*(1), 106–118. https://doi.org/10.1016/j.stem.2010.12.003

Lawson, N. D., Vogel, A. M., & Weinstein, B. M. (2002). Sonic hedgehog and vascular endothelial growth factor act upstream of the Notch pathway during arterial endothelial differentiation. *Developmental Cell*, *3*(1), 127–136. https://doi.org/10.1016/S1534-5807(02)00198-3

Lee, J. H., Protze, S. I., Laksman, Z., Backx, P. H., & Keller, G. M. (2017). Human Pluripotent Stem Cell-Derived Atrial and Ventricular Cardiomyocytes Develop from Distinct Mesoderm Populations. *Cell Stem Cell*, *21*(2), 179-194.e4. https://doi.org/10.1016/j.stem.2017.07.003

Lescroart, F., Chabab, S., Lin, X., Rulands, S., Paulissen, C., Rodolosse, A., Auer, H., Achouri, Y., Dubois, C., Bondue, A., Simons, B. D., & Blanpain, C. (2014). Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. *Nature Cell Biology*, *16*(9), 829–840. https://doi.org/10.1038/ncb3024

Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sànchez-Dànes, A., Moignard, V., Dubois, C., Paulissen, C., Kinston, S., Göttgens, B., & Blanpain, C. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science*, *359*(6380), 1177–1181. https://doi.org/10.1126/science.aao4174

Li, G., Xu, A., Sim, S., Priest, J. R., Tian, X., Khan, T., Quertermous, T., Zhou, B., Tsao, P. S., Quake, S. R., & Wu, S. M. (2016). Transcriptomic Profiling Maps Anatomically Patterned Subpopulations among Single Embryonic Cardiac Cells. *Developmental Cell*, *39*(4), 491–507. https://doi.org/10.1016/j.devcel.2016.10.014

Li, X., Martinez-Fernandez, A., Hartjes, K. A., Kocher, J. P. A., Olson, T. M., Terzic, A., & Nelson, T. J. (2014). Transcriptional atlas of cardiogenesis maps congenital heart disease interactome. *Physiological Genomics*, *46*(13), 482–495. https://doi.org/10.1152/physiolgenomics.00015.2014

Lian, X., Hsiao, C., Wilson, G., Zhu, K., Hazeltine, L. B., Azarin, S. M., Raval, K. K., Zhang, J., Kamp, T. J., & Palecek, S. P. (2012). Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(27). https://doi.org/10.1073/pnas.1200250109

Lian, X., Xu, J., Li, J., & Chien, K. R. (2014). Next-generation models of human cardiogenesis via genome editing. *Cold Spring Harbor Perspectives in Medicine*, *4*(12), a013920. https://doi.org/10.1101/cshperspect.a013920

Lian, X., Zhang, J., Azarin, S. M., Zhu, K., Hazeltine, L. B., Bao, X., Hsiao, C., Kamp, T. J., & Palecek, S. P. (2013). Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. *Nature Protocols*, *8*(1), 162–175. https://doi.org/10.1038/nprot.2012.150

Lin, S., Lin, Y., Nery, J. R., Urich, M. A., Breschi, A., Davis, C. A., Dobin, A., Zaleski, C., Beer, M. A., Chapman, W. C., Gingeras, T. R., Ecker, J. R., & Snyder, M. P. (2014). Comparison of the transcriptional landscapes between human and mouse tissues. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(48), 17224–17229. https://doi.org/10.1073/pnas.1413624111

Lindsay, E. A., Vitelli, F., Su, H., Morishima, M., Huynh, T., Pramparo, T., Jurecic, V., Ogunrinu, G., Sutherland, H. F., Scambler, P. J., Bradley, A., & Baldini, A. (2001). Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature*, *410*(6824), 97–101. https://doi.org/10.1038/35065105

Lints, T. J., Parsons, L. M., Hartley, L., Lyons, I., & Harvey, R. P. (1993). Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *Development (Cambridge, England)*, *119*(2), 419–431.

Lister, R., Pelizzola, M., Kida, Y. S., Hawkins, R. D., Nery, J. R., Hon, G., Antosiewicz-Bourget, J., Ogmalley, R., Castanon, R., Klugman, S., Downes, M., Yu, R., Stewart, R., Ren, B., Thomson, J. A., Evans, R. M., & Ecker, J. R. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*, *471*(7336), 68–73. https://doi.org/10.1038/nature09798

Little, C. D., Piquet, D. M., Davis, L. A., Walters, L., & Drake, C. J. (1989). Distribution of laminin, collagen type IV, collagen type I, and fibronectin in chicken cardiac jelly/basement membrane. *The Anatomical Record*, *224*(3), 417–425. https://doi.org/10.1002/ar.1092240310

Litviňuková, M., Talavera-López, C., Maatz, H., Reichart, D., Worth, C. L., Lindberg, E. L., Kanda, M., Polanski, K., Heinig, M., Lee, M., Nadelmann, E. R., Roberts, K., Tuck, L., Fasouli, E. S., DeLaughter, D. M., McDonough, B., Wakimoto, H., Gorham, J. M., Samari, S., … Teichmann, S. A. (2020). Cells of the adult human heart. *Nature*, 1–10. https://doi.org/10.1038/s41586-020-2797-4

Liu, D., Wang, X., He, D., Sun, C., He, X., Yan, L., Li, Y., Han, J. D. J., & Zheng, P. (2018). Single-cell RNA-sequencing reveals the existence of naive and primed pluripotency in pre-implantation rhesus monkey embryos. *Genome Research*, *28*(10), 1481–1493. https://doi.org/10.1101/gr.233437.117

Liu, X., Chen, W., Li, W., Li, Y., Priest, J. R., Zhou, B., Wang, J., & Zhou, Z. (2019). Single-Cell RNA-Seq of the Developing Cardiac Outflow Tract Reveals Convergent Development of the Vascular Smooth Muscle Cells. *Cell Reports*, *28*(5), 1346-1361.e4. https://doi.org/10.1016/j.celrep.2019.06.092

Liu, Y., Chen, S., Zühlke, L., Black, G. C., Choy, M. K., Li, N., & Keavney, B. D. (2019). Global birth prevalence of congenital heart defects 1970-2017: Updated systematic review and meta-analysis of 260 studies. *International Journal of Epidemiology*, *48*(2), 455–463. https://doi.org/10.1093/ije/dyz009

Lough, J., & Sugi, Y. (2000). Endoderm and heart development. *Dev Dyn*, 217(4):327-42. doi: 10.1002/(SICI)1097-0177(200004)217:4<327::AID-DVDY1>3.0.CO;2-K

Ma, Q., Zhou, B., & Pu, W. T. (2008). Reassessment of Isl1 and Nkx2-5 cardiac fate maps using a Gata4-based reporter of Cre activity. *Developmental Biology*, *323*(1), 98–104. https://doi.org/10.1016/j.ydbio.2008.08.013

Macatee, T. L., Hammond, B. P., Arenkiel, B. R., Francis, L., Frank, D. U., & Moon, A. M. (2003). Ablation of specific expression domains reveals discrete functions of ectoderm- and endoderm-derived FGF8 during cardiovascular and pharyngeal development. *Development (Cambridge, England)*, *130*(25), 6361–6374. https://doi.org/10.1242/dev.00850

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science*, *339*(6121), 823–826. https://doi.org/10.1126/science.1232033

Marco, E., Karp, R. L., Guo, G., Robson, P., Hart, A. H., Trippa, L., & Yuan, G.-C. (2014). Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(52), E5643-50. https://doi.org/10.1073/pnas.1408993111

Marelli, A. J., Ionescu-Ittu, R., Mackie, A. S., Guo, L., Dendukuri, N., & Kaouache, M. (2014). Lifetime prevalence of congenital heart disease in the general population from 2000 to 2010. *Circulation*, *130*(9), 749–756. https://doi.org/10.1161/CIRCULATIONAHA.113.008396

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(12 II), 7634–7638. https://doi.org/10.1073/pnas.78.12.7634

Matsa, E., & Denning, C. (2012). In vitro uses of human pluripotent stem cell-derived cardiomyocytes. *Journal of Cardiovascular Translational Research*, *5*(5), 581–592. https://doi.org/10.1007/s12265-012-9376-5

McDaniell, R., Warthen, D. M., Sanchez-Lara, P. A., Pai, A., Krantz, I. D., Piccoli, D. A., & Spinner, N. B. (2006). NOTCH2 mutations cause Alagille syndrome, a heterogeneous disorder of the notch signaling pathway. *American Journal of Human Genetics*, *79*(1), 169–173. https://doi.org/10.1086/505332

McFadden, D. G., Barbosa, A. C., Richardson, J. A., Schneider, M. D., Srivastava, D., & Olson, E. N. (2005). The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. *Development*, *132*(1), 189–201. https://doi.org/10.1242/dev.01562

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Meilhac, S. M., & Buckingham, M. E. (2018). The deployment of cell lineages that form the mammalian heart. In *Nature Reviews Cardiology* (Vol. 15, Issue 11, pp. 705–724). Nature Publishing Group. https://doi.org/10.1038/s41569-018-0086-9

Meilhac, S. M., Esner, M., Kelly, R. G., Nicolas, J. F., & Buckingham, M. E. (2004). The clonal origin of myocardial cells in different regions of the embryonic mouse heart. *Developmental Cell*, *6*(5), 685–698. https://doi.org/10.1016/S1534-5807(04)00133-9

Melnikov, A., Zhang, X., Rogov, P., Wang, L., & Mikkelsen, T. S. (2014). Massively parallel reporter assays in cultured mammalian cells. *Journal of Visualized Experiments*, *90*. https://doi.org/10.3791/51719

Merscher, S., Funke, B., Epstein, J. A., Heyer, J., Puech, A., Lu, M. M., Xavier, R. J., Demay, M. B., Russell, R. G., Factor, S., Tokooya, K., Jore, B. S., Lopez, M., Pandita, R. K., Lia, M., Carrion, D., Xu, H., Schorle, H., Kobler, J. B., … Kucherlapati, R. (2001). TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell*, *104*(4), 619–629. https://doi.org/10.1016/S0092-8674(01)00247-1

Mikawa, T., Borisov, A., Brown, A. M. C., & Fischman, D. A. (1992). Clonal analysis of cardiac morphogenesis in the chicken embryo using a replication-defective retrovirus: I. Formation of the ventricular myocardium. *Developmental Dynamics*, *193*(1), 11–23. https://doi.org/10.1002/aja.1001930104

Mikawa, T., & Gourdie, R. G. (1996). Pericardial mesoderm generates a population of coronary smooth muscle cells migrating into the heart along with ingrowth of the epicardial organ. *Developmental Biology*, *174*(2), 221–232. https://doi.org/10.1006/dbio.1996.0068

Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., Dulay, G. P., Hua, K. L., Ankoudinova, I., Cost, G. J., Urnov, F. D., Zhang, H. S., Holmes, M. C., Zhang, L., Gregory, P. D., & Rebar, E. J. (2011). A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology*, *29*(2), 143–150. https://doi.org/10.1038/nbt.1755

Mitchell E., Gilbert M., Loomes K.M. (2018). Alagille Syndrome. *Clin Liver Dis*, 22(4):625-641. doi: 10.1016/j.cld.2018.06.001. Epub 2018 Aug 22.

Mjaatvedt, C. H., Nakaoka, T., Moreno-Rodriguez, R., Norris, R. A., Kern, M. J., Eisenberg, C. A., Turner, D., & Markwald, R. R. (2001). The outflow tract of the heart is recruited from a novel heart-forming field. *Developmental Biology*, *238*(1), 97–109. https://doi.org/10.1006/dbio.2001.0409

Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., Marioni, J. C., & Reik, W. (2017). Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. *Cell Reports*, *20*(5), 1215–1228. https://doi.org/10.1016/j.celrep.2017.07.009

Mononen, M. M., Leung, C. Y., Xu, J., & Chien, K. R. (2020). Trajectory mapping of human embryonic stem cell cardiogenesis reveals lineage branch points and an ISL1 progenitor-derived cardiac fibroblast lineage. *Stem Cells*. https://doi.org/10.1002/stem.3236

Moon, A. (2008). Mouse models of congenital cardiovascular disease. *Current Topics in Developmental Biology*, *84*, 171–248. https://doi.org/10.1016/S0070-2153(08)00604-2

Moore-Morris, T., Guimarães-Camboa, N., Yutzey, K. E., Pucéat, M., & Evans, S. M. (2015). Cardiac fibroblasts: from development to heart failure. *Journal of Molecular Medicine (Berlin, Germany)*, *93*(8), 823–830. https://doi.org/10.1007/s00109-015-1314-y

Moorman, A. F. M., Christoffels, V. M., Anderson, R. H., & van den Hoff, M. J. B. (2007). The heart-forming fields: one or multiple? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1484), 1257–1265. https://doi.org/10.1098/rstb.2007.2113

Moretti, A., Caron, L., Nakano, A., Lam, J. T., Bernshausen, A., Chen, Y., Qyang, Y., Bu, L., Sasaki, M., Martin-Puig, S., Sun, Y., Evans, S. M., Laugwitz, K. L., & Chien, K. R. (2006). Multipotent Embryonic Isl1+ Progenitor Cells Lead to Cardiac, Smooth Muscle, and Endothelial Cell Diversification. *Cell*, *127*(6), 1151–1165. https://doi.org/10.1016/j.cell.2006.10.029

Morgenthau, A., & Frishman, W. H. (2018). Genetic Origins of Tetralogy of Fallot. In *Cardiology in Review* (Vol. 26, Issue 2, pp. 86–92). Lippincott Williams and Wilkins. https://doi.org/10.1097/CRD.0000000000000170

Motoike, T., Markham, D. W., Rossant, J., & Sato, T. N. (2003). Evidence for novel fate of Flk1+ progenitor: contribution to muscle lineage. *Genesis (New York, N.Y. : 2000)*, *35*(3), 153–159. https://doi.org/10.1002/gene.10175

Mummery, C. L., Zhang, J., Ng, E. S., Elliott, D. A., Elefanty, A. G., & Kamp, T. J. (2012). Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview. *Circulation Research*, *111*(3), 344–358. https://doi.org/10.1161/CIRCRESAHA.110.227512

Murray, G. (1948). Closure of Defects in Cardiac Septa. *Annals of Surgery*, *128*(4), 843–852. http://www.ncbi.nlm.nih.gov/pubmed/17859236

Naito, A. T., Shiojima, I., Akazawa, H., Hidaka, K., Morisaki, T., Kikuchi, A., & Komuro, I. (2006). Developmental stage-specific biphasic roles of Wnt/beta-catenin signaling in cardiomyogenesis and hematopoiesis. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(52), 19812–19817. https://doi.org/10.1073/pnas.0605768103

Nakajima, Y. (2010). Second lineage of heart forming region provides new understanding of conotruncal heart defects. *Congenital Anomalies*, *50*(1), 8–14. https://doi.org/10.1111/j.1741-4520.2009.00267.x

Nanba, D., & Higashiyama, S. (2004). Dual intracellular signaling by proteolytic cleavage of membrane-anchored heparin-binding EGF-like growth factor. In *Cytokine and Growth Factor Reviews* (Vol. 15, Issue 1, pp. 13–19). Elsevier BV. https://doi.org/10.1016/j.cytogfr.2003.10.002

Nascone, N., & Mercola, M. (1995). An inductive role for the endoderm in Xenopus cardiogenesis. *Development (Cambridge, England)*, *121*(2), 515–523.

Neeb, Z., Lajiness, J. D., Bolanis, E., & Conway, S. J. (2013). Cardiac outflow tract anomalies. In *Wiley Interdisciplinary Reviews: Developmental Biology* (Vol. 2, Issue 4, pp. 499–530). Wiley Interdiscip Rev Dev Biol. https://doi.org/10.1002/wdev.98

Nees, S. N., & Chung, W. K. (2020). The genetics of isolated congenital heart disease. *American Journal of Medical Genetics, Part C: Seminars in Medical Genetics*, *184*(1), 97–106. https://doi.org/10.1002/ajmg.c.31763

Niakan, K. K., & Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. *Developmental Biology*, *375*(1), 54–64. https://doi.org/10.1016/j.ydbio.2012.12.008

Niessen, K., & Karsan, A. (2008). Notch signaling in cardiac development. In *Circulation Research* (Vol. 102, Issue 10, pp. 1169–1181). Circ Res. https://doi.org/10.1161/CIRCRESAHA.108.174318

Osafune, K., Caron, L., Borowiak, M., Martinez, R. J., Fitz-Gerald, C. S., Sato, Y., Cowan, C. A., Chien, K. R., & Melton, D. A. (2008). Marked differences in differentiation propensity among human embryonic stem cell lines. *Nature Biotechnology*, *26*(3), 313–315. https://doi.org/10.1038/nbt1383

Page, D. J., Miossec, M. J., Williams, S. G., Monaghan, R. M., Fotiou, E., Cordell, H. J., Sutcliffe, L., Topf, A., Bourgey, M., Bourque, G., Eveleigh, R., Dunwoodie, S. L., Winlaw, D. S., Bhattacharya, S., Breckpot, J., Devriendt, K., Gewillig, M., Brook, J. D., Setchfield, K. J., … Keavney, B. D. (2019). Whole Exome Sequencing Reveals the Major Genetic Contributors to Nonsyndromic Tetralogy of Fallot. *Circulation Research*, *124*(4), 553–563. https://doi.org/10.1161/CIRCRESAHA.118.313250

Paige, S. L., Osugi, T., Afanasiev, O. K., Pabon, L., Reinecke, H., & Murry, C. E. (2010). Endogenous Wnt/beta-catenin signaling is required for cardiac differentiation in human embryonic stem cells. *PloS One*, *5*(6), e11134. https://doi.org/10.1371/journal.pone.0011134

Paik, D. T., Cho, S., Tian, L., Chang, H. Y., & Wu, J. C. (2020). Single-cell RNA sequencing in cardiovascular development, disease and medicine. *Nature Reviews Cardiology*. https://doi.org/10.1038/s41569-020-0359-y

Palpant, N. J., Pabon, L., Friedman, C. E., Roberts, M., Hadland, B., Zaunbrecher, R. J., Bernstein, I., Zheng, Y., & Murry, C. E. (2017). Generating high-purity cardiac and endothelial derivatives from patterned mesoderm using human pluripotent stem cells. *Nature Protocols*, *12*(1), 15–31. https://doi.org/10.1038/nprot.2016.153

Pandur, P., Sirbu, I. O., Kühl, S. J., Philipp, M., & Kühl, M. (2013). Islet1-expressing cardiac progenitor cells: a comparison across species. *Development Genes and Evolution*, *223*(1–2), 117–129. https://doi.org/10.1007/s00427-012-0400-1

Parikh, S. S., Blackwell, D. J., Gomez-Hurtado, N., Frisk, M., Wang, L., Kim, K., Dahl, C. P., Fiane, A., Tønnessen, T., Kryshtal, D. O., Louch, W. E., & Knollmann, B. C. (2017). Thyroid and Glucocorticoid Hormones Promote Functional T-Tubule Development in Human-Induced Pluripotent Stem Cell-Derived Cardiomyocytes. *Circulation Research*, *121*(12), 1323–1330. https://doi.org/10.1161/CIRCRESAHA.117.311920

Park, E. J., Ogden, L. A., Talbot, A., Evans, S., Cai, C.-L., Black, B. L., Frank, D. U., & Moon, A. M. (2006). Required, tissue-specific roles for Fgf8 in outflow tract formation and remodeling. *Development (Cambridge, England)*, *133*(12), 2419–2433. https://doi.org/10.1242/dev.02367

Parrotta, E. I., Scalise, S., Scaramuzzino, L., & Cuda, G. (2019). Stem cells: The game changers of human cardiac disease modelling and regenerative medicine. In *International Journal of Molecular Sciences* (Vol. 20, Issue 22). MDPI AG. https://doi.org/10.3390/ijms20225760

Patsch, C., Challet-Meylan, L., Thoma, E. C., Urich, E., Heckel, T., O'Sullivan, J. F., Grainger, S. J., Kapp, F. G., Sun, L., Christensen, K., Xia, Y., Florido, M. H. C., He, W., Pan, W., Prummer, M., Warren, C. R., Jakob-Roetne, R., Certa, U., Jagasia, R., … Cowan, C. A. (2015). Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. *Nature Cell Biology*, *17*(8), 994–1003. https://doi.org/10.1038/ncb3205

Pérez-Pomares, J. M., Macías, D., García-Garrido, L., & Muñoz-Chápuli, R. (1997). Contribution of the primitive epicardium to the subepicardial mesenchyme in hamster and chick embryos. *Developmental Dynamics*, *210*(2), 96–105. https://doi.org/10.1002/(SICI)1097-0177(199710)210:2<96::AID-AJA3>3.0.CO;2-4

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S. P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., & Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell*, *165*(4), 1012–1026. https://doi.org/10.1016/j.cell.2016.03.023

Pezzolla, D., López-Beas, J., Lachaud, C. C., Domínguez-Rodríguez, A., Smani, T., Hmadcha, A., & Soria, B. (2015). Resveratrol ameliorates the maturation process of β-cell-like cells obtained from an optimized differentiation protocol of human embryonic stem cells. *PloS One*, *10*(3), e0119904. https://doi.org/10.1371/journal.pone.0119904

Picelli, S., Faridani, O. R., Björklund, Å. K., Winberg, G., Sagasser, S., & Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols*, *9*(1), 171–181. https://doi.org/10.1038/nprot.2014.006

Pierpont, M. E., Brueckner, M., Chung, W. K., Garg, V., Lacro, R. V., McGuire, A. L., Mital, S., Priest, J. R., Pu, W. T., Roberts, A., Ware, S. M., Gelb, B. D., & Russell, M. W. (2018). Genetic Basis for Congenital Heart Disease: Revisited: A Scientific Statement from the American Heart Association. In *Circulation* (Vol. 138, Issue 21, pp. e653–e711). Lippincott Williams and Wilkins. https://doi.org/10.1161/CIR.0000000000000606

Poelmann, R. E., Gittenberger-de Groot, A. C., Mentink, M. M. T., Bokenkamp, R., & Hogers, B. (1993). Development of the cardiac coronary vascular endothelium, studied with antiendothelial antibodies, in chicken-quail chimeras. *Circulation Research*, *73*(3), 559–568. https://doi.org/10.1161/01.RES.73.3.559

Polo, J. M., Liu, S., Figueroa, M. E., Kulalert, W., Eminli, S., Tan, K. Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., Natesan, S., Wagers, A. J., Melnick, A., Evans, T., & Hochedlinger, K. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Biotechnology*, *28*(8), 848–855. https://doi.org/10.1038/nbt.1667

Pölönen, P., Jawahar Deen, A., Leinonen, H. M., Jyrkkänen, H.-K., Kuosmanen, S., Mononen, M., Jain, A., Tuomainen, T., Pasonen-Seppänen, S., Hartikainen, J. M., Mannermaa, A., Nykter, M., Tavi, P., Johansen, T., Heinäniemi, M., & Levonen, A.-L. (2019). Nrf2 and SQSTM1/p62 jointly contribute to mesenchymal transition and invasion in glioblastoma. *Oncogene*, *38*(50). https://doi.org/10.1038/s41388-019-0956-6

Potts, J. D., & Runyan, R. B. (1989). Epithelial-mesenchymal cell transformation in the embryonic heart can be mediated, in part, by transforming growth factor β. *Developmental Biology*, *134*(2), 392–401. https://doi.org/10.1016/0012-1606(89)90111-5

Preuss, C., Capredon, M., Wünnemann, F., Chetaille, P., Prince, A., Godard, B., Leclerc, S., Sobreira, N., Ling, H., Awadalla, P., Thibeault, M., Khairy, P., Samuels, M. E., & Andelfinger, G. (2016). Family Based Whole Exome Sequencing Reveals the Multifaceted Role of Notch Signaling in Congenital Heart Disease. *PLoS Genetics*, *12*(10), e1006335. https://doi.org/10.1371/journal.pgen.1006335

Protze, S. I., Liu, J., Nussinovitch, U., Ohana, L., Backx, P. H., Gepstein, L., & Keller, G. M. (2017). Sinoatrial node cardiomyocytes derived from human pluripotent cells function as a biological pacemaker. *Nature Biotechnology*, *35*(1), 56–68. https://doi.org/10.1038/nbt.3745

Protze, S. I., Lee, J. H., & Keller, G. M. (2019). Human Pluripotent Stem Cell-Derived Cardiovascular Cells: From Developmental Biology to Therapeutic Applications. In *Cell Stem Cell* (Vol. 25, Issue 3, pp. 311–327). Cell Press. https://doi.org/10.1016/j.stem.2019.07.010

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., & Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods*, *14*(10), 979–982. https://doi.org/10.1038/nmeth.4402

Quaranta, R., Fell, J., Rü Hle, F., Rao, J., Piccini, I., Araú Zo-Bravo, M. J., Verkerk, A. O., Stoll, M., & Greber, B. (2018). Revised roles of ISL1 in a hES cell-based model of human heart chamber specification. *Elife*, 7:e31706. doi: 10.7554/eLife.31706

Rammeloo, L. A. J., DeRuiter, M. C., van den Akker, N. M., Wisse, L. J., & Gittenberger-De Groot, A. C. (2014). Development of major aorto-pulmonary collateral arteries in vegf120/120 isoform mouse embryos with tetralogy of Fallot. *Pediatric Cardiology*, *36*(1), 89–95. https://doi.org/10.1007/s00246-014-0969-4

Rana, M. S., Théveniau-Ruissy, M., De Bono, C., Mesbah, K., Francou, A., Rammah, M., Domínguez, J. N., Roux, M., Laforest, B., Anderson, R. H., Mohun, T., Zaffran, S., Christoffels, V. M., & Kelly, R. G. (2014). Tbx1 coordinates addition of posterior second heart field progenitor cells to the arterial and venous poles of the heart. *Circulation Research*, *115*(9), 790–799. https://doi.org/10.1161/CIRCRESAHA.115.305020

Reuter, M. S., Jobling, R., Chaturvedi, R. R., Manshaei, R., Costain, G., Heung, T., Curtis, M., Hosseini, S. M., Liston, E., Lowther, C., Oechslin, E., Sticht, H., Thiruvahindrapuram, B., Mil, S. van, Wald, R. M., Walker, S., Marshall, C. R., Silversides, C. K., Scherer, S. W., … Bassett, A. S. (2019). Haploinsufficiency of vascular endothelial growth factor related signaling genes is associated with tetralogy of Fallot. *Genetics in Medicine*, *21*(4), 1001–1007. https://doi.org/10.1038/s41436-018-0260-9

Reyes, A. P., & Lanner, F. (2017). Towards a CRISPR view of early human development: Applications, limitations and ethical concerns of genome editing in human embryos. *Development (Cambridge)*, *144*(1), 3–7. https://doi.org/10.1242/dev.139683

Richter, F., Morton, S. U., Kim, S. W., Kitaygorodsky, A., Wasson, L. K., Chen, K. M., Zhou, J., Qi, H., Patel, N., DePalma, S. R., Parfenov, M., Homsy, J., Gorham, J. M., Manheimer, K. B., Velinder, M., Farrell, A., Marth, G., Schadt, E. E., Kaltman, J. R., … Gelb, B. D. (2020). Genomic analyses implicate noncoding de novo variants in congenital heart disease. *Nature Genetics*. https://doi.org/10.1038/s41588-020-0652-z

Robertson, C., Tran, D. D., & George, S. C. (2013). Concise review: Maturation phases of human pluripotent stem cell-derived cardiomyocytes. In *Stem Cells* (Vol. 31, Issue 5, pp. 829–837). Stem Cells. https://doi.org/10.1002/stem.1331

Rokicki, W., Leung, M., Kohut, J., Durmała, J., & Borowicka, E. (2003). Population dependent differences of distribution of congenital heart malformations. *Wiadomosci Lekarskie (Warsaw, Poland : 1960)*, *56*(1–2), 34–39.

Ruan, H., Liao, Y., Ren, Z., Mao, L., Yao, F., Yu, P., Ye, Y., Zhang, Z., Li, S., Xu, H., Liu, J., Diao, L., Zhou, B., Han, L., & Wang, L. (2019). Single-cell reconstruction of differentiation trajectory reveals a critical role of ETS1 in human cardiac lineage commitment. *BMC Biology*, *17*(1), 89. https://doi.org/10.1186/s12915-019-0709-6

Saga, Y., Kitajima, S., & Miyagawa-Tomita, S. (2000). Mesp1 expression is the earliest sign of cardiovascular development. In *Trends in Cardiovascular Medicine* (Vol. 10, Issue 8, pp. 345–352). Trends Cardiovasc Med. https://doi.org/10.1016/S1050-1738(01)00069-X

Sahara, M., Santoro, F., Sohlmér, J., Zhou, C., Witman, N., Leung, C. Y., Mononen, M., Bylund, K., Gruber, P., & Chien, K. R. (2019). Population and Single-Cell Analysis of Human Cardiogenesis Reveals Unique LGR5 Ventricular Progenitors in Embryonic Outflow Tract. *Developmental Cell*, *48*(4). https://doi.org/10.1016/j.devcel.2019.01.005

Sala, L., Bellin, M., & Mummery, C. L. (2017). Integrating cardiomyocytes from human pluripotent stem cells in safety pharmacology: has the time come? *British Journal of Pharmacology*, *174*(21), 3749–3765. https://doi.org/10.1111/bph.13577

Schilham, M. W., Oosterwegel, M. A., Moerer, P., Ya, J., de Boer, P. A., van de Wetering, M., Verbeek, S., Lamers, W. H., Kruisbeek, A. M., Cumano, A., & Clevers, H. (1996). Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature*, *380*(6576), 711–714. https://doi.org/10.1038/380711a0

Schuldiner, M., Yanuka, O., Itskovitz-Eldor, J., Melton, D. A., & Benvenisty, N. (2000). Effects of eight growth factors on the differentiation of cells derived from human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(21), 11307–11312. https://doi.org/10.1073/pnas.97.21.11307

Schultheiss, T. M., Xydas, S., & Lassar, A. B. (1995). Induction of avian cardiac myogenesis by anterior endoderm. *Development (Cambridge, England)*, *121*(12), 4203–4214.

Schwach, V., Gomes Fernandes, M., Maas, S., Gerhardt, S., Tsonaka, R., Van Der Weerd, L., Passier, R., Mummery, C. L., Birket, M. J., & Salvatori, D. C. F. (2020). Expandable human cardiovascular progenitors from stem cells for regenerating mouse heart after myocardial infarction. *Cardiovascular Research*, *116*(3), 545–553. https://doi.org/10.1093/cvr/cvz181

Scialdone, A., Tanaka, Y., Jawaid, W., Moignard, V., Wilson, N. K., Macaulay, I. C., Marioni, J. C., & Göttgens, B. (2016). Resolving early mesoderm diversification through single-cell expression profiling. *Nature*, *535*(7611), 289–293. https://doi.org/10.1038/nature18633

Shabana, N., Shahid, S. U., & Irfan, U. (2020). Genetic Contribution to Congenital Heart Disease (CHD). In *Pediatric Cardiology* (Vol. 41, Issue 1, pp. 12–23). Springer. https://doi.org/10.1007/s00246-019-02271-4

Shapiro, E., Biezuner, T., & Linnarsson, S. (2013). Single-cell sequencing-based technologies will revolutionize whole-organism science. In *Nature Reviews Genetics* (Vol. 14, Issue 9, pp. 618–630). https://doi.org/10.1038/nrg3542

Sharkey, A. M., & Sharma, A. (2012). Tetralogy of Fallot: anatomic variants and their impact on surgical management. *Seminars in Cardiothoracic and Vascular Anesthesia*, *16*(2), 88–96. https://doi.org/10.1177/1089253211434566

Sharma, A. (2020). Modeling Congenital Heart Disease Using Pluripotent Stem Cells. *Current Cardiology Reports*, *22*(8), 55. https://doi.org/10.1007/s11886-020-01316-y

Shin, J., Berg, D. A., Zhu, Y., Shin, J. Y., Song, J., Bonaguidi, M. A., Enikolopov, G., Nauen, D. W., Christian, K. M., Ming, G. L., & Song, H. (2015). Single-Cell RNA-Seq with Waterfall Reveals Molecular Cascades underlying Adult Neurogenesis. *Cell Stem Cell*, *17*(3), 360–372. https://doi.org/10.1016/j.stem.2015.07.013

Showell, C., Binder, O., & Conlon, F. L. (2004). T-box genes in early embryogenesis. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *229*(1), 201–218. https://doi.org/10.1002/dvdy.10480

Sifrim, A., Hitz, M. P., Wilsdon, A., Breckpot, J., Turki, S. H. A., Thienpont, B., McRae, J., Fitzgerald, T. W., Singh, T., Swaminathan, G. J., Prigmore, E., Rajan, D., Abdul-Khaliq, H., Banka, S., Bauer, U. M. M., Bentham, J., Berger, F., Bhattacharya, S., Bu'Lock, F., … Hurles, M. E. (2016). Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nature Genetics*, *48*(9), 1060–1065. https://doi.org/10.1038/ng.3627

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., & Kucherlapati, R. S. (1985). Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. *Nature*, *317*(6034), 230–234. https://doi.org/10.1038/317230a0

Snider, P., & Conway, S. J. (2011). Probing human cardiovascular congenital disease using transgenic mouse models. *Prog Mol Biol Transl Sci*, 100:83-110. doi: 10.1016/B978-0-12-384878-9.00003-0.

Soemedi, R., Wilson, I. J., Bentham, J., Darlay, R., Töpf, A., Zelenika, D., Cosgrove, C., Setchfield, K., Thornborough, C., Granados-Riveron, J., Blue, G. M., Breckpot, J., Hellens, S., Zwolinkski, S., Glen, E., Mamasoula, C., Rahman, T. J., Hall, D., Rauch, A., … Keavney, B. D. (2012). Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *American Journal of Human Genetics*, *91*(3), 489–501. https://doi.org/10.1016/j.ajhg.2012.08.003

Somi, S., Buffing, A. A. M., Moorman, A. F. M., & Van Den Hoff, M. J. B. (2004). Dynamic patterns of expression of BMP isoforms 2, 4, 5, 6, and 7 during chicken heart development. *The Anatomical Record. Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology*, *279*(1), 636–651. https://doi.org/10.1002/ar.a.20031

Später, D., Hansson, E. M., Zangi, L., & Chien, K. R. (2014). How to make a cardiomyocyte. *Development (Cambridge, England)*, *141*(23), 4418–4431. https://doi.org/10.1242/dev.091538

Stirparo, G. G., Boroviak, T., Guo, G., Nichols, J., Smith, A., & Bertone, P. (2018). Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development*, *145*(3), dev158501. https://doi.org/10.1242/dev.158501

Stoll, C., Dott, B., Alembik, Y., & Roth, M. P. (2015). Associated congenital anomalies among cases with Down syndrome. *European Journal of Medical Genetics*, *58*(12), 674–680. https://doi.org/10.1016/j.ejmg.2015.11.003

Stoller, J. Z., & Epstein, J. A. (2005). Cardiac neural crest. In *Seminars in Cell and Developmental Biology* (Vol. 16, Issue 6, pp. 704–715). Elsevier Ltd. https://doi.org/10.1016/j.semcdb.2005.06.004

Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E., & Dudoit, S. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*, *19*(1). https://doi.org/10.1186/s12864-018-4772-0

Sun, Y., Liang, X., Najafi, N., Cass, M., Lin, L., Cai, C. L., Chen, J., & Evans, S. M. (2007). Islet 1 is expressed in distinct cardiovascular lineages, including pacemaker and coronary vascular cells. *Developmental Biology*, *304*(1), 286–296. https://doi.org/10.1016/j.ydbio.2006.12.048

Suryawanshi, H., Clancy, R., Morozov, P., Halushka, M. K., Buyon, J. P., & Tuschl, T. (2020). Cell atlas of the foetal human heart and implications for autoimmune-mediated congenital heart block. *Cardiovascular Research*, *116*(8), 1446–1457. https://doi.org/10.1093/cvr/cvz257

Svensson, V., & Pachter, L. (2018). RNA Velocity: Molecular Kinetics from Single-Cell RNA-Seq. In *Molecular Cell* (Vol. 72, Issue 1, pp. 7–9). Cell Press. https://doi.org/10.1016/j.molcel.2018.09.026

Sylva, M., Van den Hoff, M. J. B., & Moorman, A. F. M. (2014). Development of the human heart. *American Journal of Medical Genetics, Part A*, *164*(6), 1347–1371. https://doi.org/10.1002/ajmg.a.35896

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, *131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Tam, P. P., Parameswaran, M., Kinder, S. J., & Weinberger, R. P. (1997). The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. *Development (Cambridge, England)*, *124*(9), 1631–1642.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., & Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, *6*(5), 377–382. https://doi.org/10.1038/nmeth.1315

Teratani-Ota, Y., Yamamizu, K., Piao, Y., Sharova, L., Amano, M., Yu, H., Schlessinger, D., Ko, M. S. H., & Sharov, A. A. (2016). Induction of specific neuron types by overexpression of single transcription factors. *In Vitro Cellular & Developmental Biology. Animal*, *52*(9), 961–973. https://doi.org/10.1007/s11626-016-0056-7

Tevosian, S. G., Deconinck, A. E., Tanaka, M., Schinke, M., Litovsky, S. H., Izumo, S., Fujiwara, Y., & Orkin, S. H. (2000). FOG-2, a Cofactor for GATA Transcription Factors, Is Essential for Heart Morphogenesis and Development of Coronary Vessels from Epicardium. In *Cell* (Vol. 101).

Théveniau-Ruissy, M., Dandonneau, M., Mesbah, K., Ghez, O., Mattei, M. G., Miquerol, L., & Kelly, R. G. (2008). The del22q11.2 candidate gene tbx1 controls regional outflow tract identity and coronary artery patterning. *Circulation Research*, *103*(2), 142–148. https://doi.org/10.1161/CIRCRESAHA.108.172189

Thom, T., Haase, N., Rosamond, W., Howard, V. J., Rumsfeld, J., Manolio, T., Zheng, Z. J., Flegal, K., O'Donnell, C., Kittner, S., Lloyd-Jones, D., Goff, D. C., & Hong, Y. (2006). Heart disease and stroke statistics - 2006 Update: A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. In *Circulation* (Vol. 113, Issue 6). Circulation. https://doi.org/10.1161/CIRCULATIONAHA.105.171600

Thomas, K. R., & Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell*, *51*(3), 503–512. https://doi.org/10.1016/0092-8674(87)90646-5

Thomson, J. A. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, *282*(5391), 1145–1147. https://doi.org/10.1126/science.282.5391.1145

Töpf, A., Griffin, H. R., Glen, E., Soemedi, R., Brown, D. L., Hall, D., Rahman, T. J., Eloranta, J. J., Jüngst, C., Stuart, A. G., O'Sullivan, J., Keavney, B. D., & Goodship, J. A. (2014). Functionally significant, rare transcription factor variants in tetralogy of fallot. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0095453

Torres-Ruiz, R., & Rodriguez-Perales, S. (2017). CRISPR-Cas9 technology: applications and human disease modelling. *Brief Funct Genomics*, 16(1):4-12. doi: 10.1093/bfgp/elw025. Epub 2016 Jun 26

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-41695-z

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S., & Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology*, *32*. https://doi.org/10.1038/nbt.2859

Tyser, R. C. V., Ibarra-Soria, X., McDole, K., A. Jayaram, S., Godwin, J., van den Brand, T. A. H., Miranda, A. M. A., Scialdone, A., Keller, P. J., Marioni, J. C., & Srinivas, S. (2021). Characterization of a common progenitor pool of the epicardium and myocardium. *Science*, eabb2986. https://doi.org/10.1126/science.abb2986

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. In *Nature Reviews Genetics* (Vol. 11, Issue 9, pp. 636–646). https://doi.org/10.1038/nrg2842

van den Akker, N. M. S., Caolo, V., & Molin, D. G. M. (2012). Cellular decisions in cardiac outflow tract and coronary development: An act by VEGF and NOTCH. In *Differentiation* (Vol. 84, Issue 1, pp. 62–78). Differentiation. https://doi.org/10.1016/j.diff.2012.04.002

Van Den Akker, N. M. S., Molin, D. G. M., Peters, P. P. W. M., Maas, S., Wisse, L. J., Van Brempt, R., Van Munsteren, C. J., Bartelings, M. M., Poelmann, R. E., Carmeliet, P., & Gittenberger-De Groot, A. C. (2007). Tetralogy of fallot and alterations in vascular endothelial growth factor-A signaling and notch signaling in mouse embryos solely expressing the VEGF120 isoform. *Circulation Research*, *100*(6), 842–849. https://doi.org/10.1161/01.RES.0000261656.04773.39

van den Berg, C. W., Okawa, S., Chuva de Sousa Lopes, S. M., van Iperen, L., Passier, R., Braam, S. R., Tertoolen, L. G., del Sol, A., Davis, R. P., & Mummery, C. L. (2015). Transcriptome of human foetal heart compared with cardiomyocytes from pluripotent stem cells. *Development (Cambridge, England)*, *142*(18), 3231–3238. https://doi.org/10.1242/dev.123810

Van Den Berg, G., Abu-Issa, R., De Boer, B. A., Hutson, M. R., De Boer, P. A. J., Soufan, A. T., Ruijter, J. M., Kirby, M. L., Van Den Hoff, M. J. B., & Moorman, A. F. M. (2009). A caudal proliferating growth center contributes to both poles of the forming heart tube. *Circulation Research*, *104*(2), 179–188. https://doi.org/10.1161/CIRCRESAHA.108.185843

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V, Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*(1110), 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

Van Der Linde, D., Konings, E. E. M., Slager, M. A., Witsenburg, M., Helbing, W. A., Takkenberg, J. J. M., & Roos-Hesselink, J. W. (2011). Birth prevalence of congenital heart disease worldwide: A systematic review and meta-analysis. In *Journal of the American College of Cardiology* (Vol. 58, Issue 21, pp. 2241–2247). Elsevier USA. https://doi.org/10.1016/j.jacc.2011.08.025

van Praagh, R., van Praagh, S., Nebesar, R. A., Muster, A. J., Sinha, S. N., & Paul, M. H. (1970). Tetralogy of fallot: Underdevelopment of the pulmonary infundibulum and its sequelae. *The American Journal of Cardiology*, *26*(1), 25–33. https://doi.org/10.1016/0002-9149(70)90754-X

Villa-Diaz, L. G., Ross, A. M., Lahann, J., & Krebsbach, P. H. (2013). Concise review: The evolution of human pluripotent stem cell culture: From feeder cells to synthetic coatings. In *Stem Cells* (Vol. 31, Issue 1, pp. 1–7). Stem Cells. https://doi.org/10.1002/stem.1260

Vincent, S. D., & Buckingham, M. E. (2010). How to make a heart: the origin and regulation of cardiac progenitor cells. *Current Topics in Developmental Biology*, *90*, 1–41. https://doi.org/10.1016/S0070-2153(10)90001-X

Waardenberg, A. J., Ramialison, M., Bouvere, R., & Harvey, R. P. (2014). Genetic networks governing heart development. *Cold Spring Harbor Perspectives in Medicine*, *4*(11), a013839. https://doi.org/10.1101/cshperspect.a013839

Waldo, K. L., Kumiski, D. H., Wallis, K. T., Stadt, H. A., Hutson, M. R., Platt, D. H., & Kirby, M. L. (2001). Conotruncal myocardium arises from a secondary heart field. *Development*, *128*(16), 3179–3188. https://pubmed-ncbi-nlm-nih-gov.proxy.kib.ki.se/11688566/

Waldo, K. L., Hutson, M. R., Stadt, H. A., Zdanowicz, M., Zdanowicz, J., & Kirby, M. L. (2005). Cardiac neural crest is necessary for normal addition of the myocardium to the arterial pole from the secondary heart field. *Developmental Biology*, *281*(1), 66–77. https://doi.org/10.1016/j.ydbio.2005.02.011

Wang, L., Yu, P., Zhou, B., Song, J., Li, Z., Zhang, M., Guo, G., Wang, Y., Chen, X., Han, L., & Hu, S. (2020). Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nature Cell Biology*, *22*(1), 108–119. https://doi.org/10.1038/s41556-019-0446-7

Ward, C., Stadt, H., Hutson, M., & Kirby, M. L. (2005). Ablation of the secondary heart field leads to tetralogy of Fallot and pulmonary atresia. *Developmental Biology*, *284*(1), 72–83. https://doi.org/10.1016/j.ydbio.2005.05.003

Ward, L. D., & Kellis, M. (2012). Interpreting noncoding genetic variation in complex traits and human disease. *Nature Biotechnology*, *30*(11), 1095–1106. https://doi.org/10.1038/nbt.2422

Ware, J. S., Samocha, K. E., Homsy, J., & Daly, M. J. (2015). Interpreting de novo Variation in Human Disease Using denovolyzeR. *Current Protocols in Human Genetics*, *87*(1), 7.25.1-7.25.15. https://doi.org/10.1002/0471142905.hg0725s87

Watanabe, Y., Miyagawa-Tomita, S., Vincent, S. D., Kelly, R. G., Moon, A. M., & Buckingham, M. E. (2010). Role of mesodermal FGF8 and FGF10 overlaps in the development of the arterial pole of the heart and pharyngeal arch arteries. *Circulation Research*, *106*(3), 495–503. https://doi.org/10.1161/CIRCRESAHA.109.201665

West, R. C., Ming, H., Logsdon, D. M., Sun, J., Rajput, S. K., Kile, R. A., Schoolcraft, W. B., Michael Roberts, R., Krisher, R. L., Jiang, Z., & Yuan, Y. (2019). Dynamics of trophoblast differentiation in peri-implantation–stage human embryos. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(45), 22635–22644. https://doi.org/10.1073/pnas.1911362116

Williams, K., Carson, J., & Lo, C. (2019). Genetics of congenital heart disease. In *Biomolecules* (Vol. 9, Issue 12). https://doi.org/10.3390/biom9120879

Wimalasundera, R. C., & Gardiner, H. M. (2004). Congenital heart disease and aneuploidy. In *Prenatal Diagnosis* (Vol. 24, Issue 13, pp. 1116–1122). https://doi.org/10.1002/pd.1068

Wiputra, H., Chen, C. K., Talbi, E., Lim, G. L., Soomar, S. M., Biswas, A., Mattar, C. N. Z., Bark, D., Leo, H. L., & Yap, C. H. (2018). Human fetal hearts with tetralogy of fallot have altered fluid dynamics and forces. *American Journal of Physiology - Heart and Circulatory Physiology*, *315*(6), H1649–H1659. https://doi.org/10.1152/ajpheart.00235.2018

Witman, N., Zhou, C., Grote Beverborg, N., Sahara, M., & Chien, K. R. (2020). Cardiac progenitors and paracrine mediators in cardiogenesis and heart regeneration. In *Seminars in Cell and Developmental Biology* (Vol. 100, pp. 29–51). Elsevier Ltd. https://doi.org/10.1016/j.semcdb.2019.10.011

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, *2*(1–3), 37–52. https://doi.org/10.1016/0169-7439(87)80084-9

Wu, B., Wang, Y., Lui, W., Langworthy, M., Tompkins, K. L., Hatzopoulos, A. K., Baldwin, H. S., & Zhou, B. (2011). Nfatc1 coordinates valve endocardial cell lineage development required for heart valve formation. *Circulation Research*, *109*(2), 183–192. https://doi.org/10.1161/CIRCRESAHA.111.245035

Xiang, L., Yin, Y., Zheng, Y., Ma, Y., Li, Y., Zhao, Z., Guo, J., Ai, Z., Niu, Y., Duan, K., He, J., Ren, S., Wu, D., Bai, Y., Shang, Z., Dai, X., Ji, W., & Li, T. (2020). A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature*, *577*(7791), 537–542. https://doi.org/10.1038/s41586-019-1875-y

Xie, S., Cooley, A., Armendariz, D., Zhou, P., & Hon, G. C. (2018). Frequent sgRNA-barcode recombination in single-cell perturbation assays. *PLoS ONE*, *13*(6). https://doi.org/10.1371/journal.pone.0198635

Xie, Y., Ma, A., Wang, B., Peng, R., Jing, Y., Wang, D., Finnell, R. H., Qiao, B., Wang, Y., Wang, H., & Zheng, Y. (2019). Rare mutations of ADAM17 from TOFs induce hypertrophy in human embryonic stem cell-derived cardiomyocytes via HB-EGF signaling. *Clinical Science*, *133*(2), 225–238. https://doi.org/10.1042/CS20180842

Xiong, H., Luo, Y., Yue, Y., Zhang, J., Ai, S., Li, X., Wang, X., Zhang, Y. L., Wei, Y., Li, H. H., Hu, X., Li, C., & He, A. (2019). Single-Cell Transcriptomics Reveals Chemotaxis-Mediated Intraorgan Crosstalk during Cardiogenesis. *Circulation Research*, *125*(4), 398–410. https://doi.org/10.1161/CIRCRESAHA.119.315243

Xu, X. Q., Graichen, R., Soo, S. Y., Balakrishnan, T., Rahmat, S. N. B., Sieh, S., Tham, S. C., Freund, C., Moore, J., Mummery, C., Colman, A., Zweigerdt, R., & Davidson, B. P. (2008). Chemically defined medium supporting cardiomyocyte differentiation of human embryonic stem cells. *Differentiation; Research in Biological Diversity*, *76*(9), 958–970. https://doi.org/10.1111/j.1432-0436.2008.00284.x

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., Huang, J., Li, M., Wu, X., Wen, L., Lao, K., Li, R., Qiao, J., & Tang, F. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nature Structural and Molecular Biology*, *20*(9), 1131–1139. https://doi.org/10.1038/nsmb.2660

Yang, L., Soonpaa, M. H., Adler, E. D., Roepke, T. K., Kattman, S. J., Kennedy, M., Henckaerts, E., Bonham, K., Abbott, G. W., Linden, R. M., Field, L. J., & Keller, G. M. (2008). Human cardiovascular progenitor cells develop from a KDR+ embryonic-stem-cell-derived population. *Nature*, *453*(7194), 524–528. https://doi.org/10.1038/nature06894

Yang, Y. P., Li, H. R., Cao, X. M., Wang, Q. X., Qiao, C. J., & Ya, J. (2013). Second heart field and the development of the outflow tract in human embryonic heart. *Development Growth and Differentiation*, *55*(3), 359–367. https://doi.org/10.1111/dgd.12050

Yang, Y. Q., Gharibeh, L., Li, R. G., Xin, Y. F., Wang, J., Liu, Z. M., Qiu, X. B., Xu, Y. J., Xu, L., Qu, X. K., Liu, X., Fang, W. Y., Huang, R. T., Xue, S., & Nemer, G. (2013). GATA4 Loss-of-Function Mutations Underlie Familial Tetralogy of Fallot. *Human Mutation*, *34*(12), 1662–1671. https://doi.org/10.1002/humu.22434

Ye, L., Ni, X., Zhao, Z. A., Lei, W., & Hu, S. (2018). The application of induced pluripotent stem cells in cardiac disease modeling and drug testing. *Journal of Cardiovascular Translational Research*, *11*(5), 366–374. https://doi.org/10.1007/s12265-018-9811-3

Yelbuz, T. M., Waldo, K. L., Kumiski, D. H., Stadt, H. A., Wolfe, R. R., Leatherbury, L., & Kirby, M. L. (2002). Shortened outflow tract leads to altered cardiac looping after neural crest ablation. *Circulation*, *106*(4), 504–510. https://doi.org/10.1161/01.CIR.0000023044.44974.8A

Ylä-Herttuala, S., Rissanen, T. T., Vajanto, I., & Hartikainen, J. (2007). Vascular endothelial growth factors: biology and current status of clinical applications in cardiovascular medicine. *Journal of the American College of Cardiology*, *49*(10), 1015–1026. https://doi.org/10.1016/j.jacc.2006.09.053

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, I. I., & Thomson, J. A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science*, *318*(5858), 1917–1920. https://doi.org/10.1126/science.1151526

Zaffran, S., & Frasch, M. (2002). Early signals in cardiac development. In *Circulation Research* (Vol. 91, Issue 6, pp. 457–469). Circ Res. https://doi.org/10.1161/01.RES.0000034152.74523.A8

Zaffran, S., Kelly, R. G., Meilhac, S. M., Buckingham, M. E., & Brown, N. A. (2004). Right ventricular myocardium derives from the anterior heart field. *Circulation Research*, *95*(3), 261–268. https://doi.org/10.1161/01.RES.0000136815.73623.BE

Zhang, J., Tao, R., Campbell, K. F., Carvalho, J. L., Ruiz, E. C., Kim, G. C., Schmuck, E. G., Raval, A. N., da Rocha, A. M., Herron, T. J., Jalife, J., Thomson, J. A., & Kamp, T. J. (2019). Functional cardiac fibroblasts derived from human pluripotent stem cells via second heart field progenitors. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-09831-5

Zhang, W., Shen, L., Deng, Z., Ding, Y., Mo, X., Xu, Z., Gao, Q., & Yi, L. (2014). Novel missense variants of ZFPM2/FOG2 identified in conotruncal heart defect patients do not impair interaction with GATA4. *PLoS ONE*, *9*(7). https://doi.org/10.1371/journal.pone.0102379

Zhao, A.-D., Qin, H., Sun, M.-L., Ma, K., & Fu, X.-B. (2020). Efficient and rapid conversion of human astrocytes and ALS mouse model spinal cord astrocytes into motor neuron-like cells by defined small molecules. *Military Medical Research*, *7*(1), 42. https://doi.org/10.1186/s40779-020-00271-7

Zhou, B., & Wang, L. (2020). Transcriptional Profiling of Single Cardiomyocytes in Health and Disease. *Curr Cardiol Rep,* 22(9):92. doi: 10.1007/s11886-020-01346-6.

Zhou, P., & Pu, W. T. (2016). Recounting Cardiac Cellular Composition. In *Circulation research* (Vol. 118, Issue 3, pp. 368–370). https://doi.org/10.1161/CIRCRESAHA.116.308139

Zhou, Q., Wang, T., Leng, L., Zheng, W., Huang, J., Fang, F., Yang, L., Chen, F., Lin, G., Wang, W. J., & Kristiansen, K. (2019). Single-cell RNA-seq reveals distinct dynamic behavior of sex chromosomes during early human embryogenesis. *Molecular Reproduction and Development*, *86*(7), 871–882. https://doi.org/10.1002/mrd.23162

Zhou, Y., Kim, J., Yuan, X., & Braun, T. (2011). Epigenetic modifications of stem cells: a paradigm for the control of cardiac progenitor cells. *Circulation Research*, *109*(9), 1067–1081. https://doi.org/10.1161/CIRCRESAHA.111.243709

Zhu, H., Kartiko, S., & Finnell, R. H. (2009). Importance of gene - Environment interactions in the etiology of selected birth defects. In *Clinical Genetics* (Vol. 75, Issue 5, pp. 409–423). https://doi.org/10.1111/j.1399-0004.2009.01174.x

Zhu, Z., & Huangfu, D. (2013). Human pluripotent stem cells: An emerging model in developmental biology. In *Development (Cambridge)* (Vol. 140, Issue 4, pp. 705–717). Company of Biologists. https://doi.org/10.1242/dev.086165

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., & Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular Cell*, *65*(4), 631-643.e4. https://doi.org/10.1016/j.molcel.2017.01.023

Zimmerman, M. S., Smith, A. G. C., Sable, C. A., Echko, M. M., Wilner, L. B., Olsen, H. E., Atalay, H. T., Awasthi, A., Bhutta, Z. A., Boucher, J. L. A., Castro, F., Cortesi, P. A., Dubey, M., Fischer, F., Hamidi, S., Hay, S. I., Hoang, C. L., Hugo-Hamman, C., Jenkins, K. J., … Kassebaum, N. J. (2020). Global, regional, and national burden of congenital heart disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Child and Adolescent Health*, *4*(3), 185–200. https://doi.org/10.1016/S2352-4642(19)30402-X